# Robust Adaptive Rate-Optimal Testing for the White Noise Hypothesis [1]

(with supplementary material)

Alain Guay[2]

Emmanuel Guerre[3]

Štěpána Lazarová[4]

This version: 2nd November 2012

---

[1]This is a revised version of a paper previously entitled 'Adaptive Rate-Optimal Detection of Small Correlation Coefficients'. We would like to thank Richard Davis, Marcelo Fernandes, Liudas Giraitis, Bruce Hansen, George Kapetanios, Remigijus Leipus, Peter Phillips and Aris Spanos for stimulating questions and suggestions. We would also like to thank participants of the Queen Mary Econometric Reading Group, Economic Seminar at York, Southampton and Warwick Universities, 2008 Oxbridge Time Series Workshop, 2008 Vienna Model Selection Workshop, 2008 North American Econometric Society Conference and 2008 European Econometric Society Conference, Journées 2008 de Statistiques de Rennes, 2009 Bristol Econometrics Worshop and 2011 Brazilian Time Series and Econometrics School. Last but not least, we would like to thank an anonymous associate editor and two referees whose feedback helped to improve the paper. All remaining errors are ours. The first author gratefully acknowledges financial support of Fonds de recherche sur la société et la culture (FQRSC) and the Social Sciences and Humanities Research Council of Canada (SSHRC). The last two authors are thankful for the financial support from the School of Economics and Finance of Queen Mary, University of London.

[2]CIRPÉE and CIREQ, Université du Québec à Montréal, e-mail: guay.alain@uqam.ca

[3]School of Economics and Finance, Queen Mary, University of London, e-mail: e.guerre@qmul.ac.uk

[4]School of Economics and Finance, Queen Mary, University of London, e-mail: s.lazarova@qmul.ac.uk

## Abstract

A new test is proposed for the weak white noise null hypothesis. The test is based on a new automatic selection of the order for a Box-Pierce (1970) test statistic or the test statistic of Hong (1996). The heteroskedasticity and autocorrelation-consistent (HAC) critical values from Lee (2007) are used, allowing for estimation of the error term. The data-driven order selection is tailored to detect a new class of alternatives with autocorrelation coefficients which can be $o(n^{-1/2})$ provided there is sufficiently many of such coefficients. A simulation experiment illustrates the good statistical properties of the test both under the weak white noise null and the alternative.

*JEL Classification*: Primary C12; Secondary C32.

*Keywords*: Weak white noise hypothesis; HAC Inference; Automatic nonparametric tests; Adaptive rate-optimality.

## 1. Introduction

Testing for white noise is important in many econometric contexts. Ignoring autocorrelation of the error terms in a linear regression model can lead to erroneous confidence intervals and tests. Correlation of residuals from an ARMA model or of the squared residuals from an ARCH model can indicate an improper choice of the order. Investigation of autocorrelation function is also a popular diagnostic tool in macroeconomics and finance, see e.g. Durlauf (1991) and Campbell, Lo and Craig MacKinlay (1997). Earliest tests of the white noise hypothesis were based on confidence intervals for autocorrelation coefficients as described by Fan and Yao (2005). See also Xiao and Wu (2011) who have recently derived the asymptotic distribution of the maximum standardized sample covariance of weak white noise, that is an stationary process which is uncorrelated but possibly dependent. A second approach was established by Grenander and Rosenblatt (1952) who extended goodness-of-fit tests such as Kolmogorov and Cramér-von Mises tests to tests of white noise hypothesis. Grenander and Rosenblatt (1952) has been refined by Durlauf (1991), Anderson (1993) and Deo (2000). Delgado, Hidalgo and Velasco (2005) have studied a modified test statistic to be used with residuals. Shao (2011a) has recently extended this setup to cover the weak white noise null hypothesis. A third approach, pioneered by Box and Pierce (1970), is based on the sum of squared sample autocorrelation coefficients up to a given order $p$. Delgado and Velasco (2012), Francq, Roy and Zakoian (2005), Kuan and Lee (2006) and Lobato (2001) have considered the weak white noise hypothesis. The case where $p$ grows with the sample size $n$ has been considered by Hong (1996) in a strong white noise setup and recently extended to the weak white noise null hypothesis by Shao (2011b) and Xiao and Wu (2011).

This paper contributes to the literature by proposing a data-driven choice $\hat{p}$ of the order $p$ used in a Box-Pierce type statistic for a test of the weak white noise null hypothesis. Under this null, $\widehat{p}$ tends to 1 in probability so that the null limit behavior of the test statistic is driven by the first-order sample autocovariance. It is shown that the test can be implemented using robust critical values of Lee (2007) who extends the work of Lobato (2001) for the case of observed variables and of Kuan and Lee (2006) for the case of residuals. The general framework of Lee (2007) includes as a specific case standardization using steep origin kernels proposed by Phillips, Sun and Jin (2006) which can improve the power of the resulting test. Under the alternative, the data-driven $\widehat{p}$ can be as large as necessary.

An appealing feature of Cramér-von Mises type of tests is the ability to detect Pitman local directional alternatives converging to the null with the parametric rate $n^{-1/2}$. This contrasts with detection results for Box-Pierce type test by Hong (1996) which is only consistent under slower rates of convergence for local alternatives defined through the spectral density function. The conclusions of Hong (1996) suggest that Cramér-von Mises tests are more powerful than Box-Pierce tests. One of the contributions of the present paper is to point out that this ranking of two types of tests is not universal and there exist classes of alternatives against which Box-Pierce tests are more powerful than Cramér-von Mises tests.

We illustrate this point using a new class of alternatives defined through the autocovariance function. The new class of alternatives formalizes the idea that small autocorrelation coefficients of magnitude $\rho_n$ can be detected provided that there are sufficiently many coefficients present at smaller lags. An important finding of the paper is that detection is still possible for very small $\rho_n$, namely for $\rho_n = o\left(n^{-1/2}\right)$. As described in Section 4, this type of alternatives includes moving average processes with a significant long term multiplier but $o\left(n^{-1/2}\right)$ impulse response coefficients. Such processes therefore correspond to a macroeconomic scenario where short term policies have no significant effects whereas long term policies may have an impact. For such alternatives, the conditional expectation of the present given the past gives $o\left(n^{-1/2}\right)$ weights to each lagged observations. Therefore this process is hard to predict since it is very close to a martingale difference process. These alternatives can be of interest in finance where arbitrage could forbid strong deviations from martingale difference.

Why such alternatives can be detected by Box-Pierce tests can be intuitively explained as follows. Let $\widehat{R}_j$ and $R_j$ be respectively the sample and population covariance at lag $j$. Following Hong (1996), Shao (2011b) and Xiao and Wu (2011), the nonrobust critical region of the Box-Pierce test of order $p_n \to \infty$ is

$$\frac{n \sum_{j=1}^{p_n} \left( \widehat{R}_j^2 / \widehat{R}_0^2 - 1 \right)}{(2p_n)^{1/2}} \geq c_\alpha, \tag{1.1}$$

where $c_\alpha$ is a standard normal critical value. Arguing as Shao (2011b, Theorem 2.2) suggests that

$$\frac{n \sum_{j=1}^{p_n} \left( \widehat{R}_j^2 / \widehat{R}_0^2 - 1 \right)}{(2p_n)^{1/2}} = \frac{n \sum_{j=1}^{p_n} R_j^2 / R_0^2}{(2p_n)^{1/2}} + O_{\mathbb{P}}(1). \tag{1.2}$$

(1.2) suggests that the Box-Pierce test is consistent provided $\left( n / (2p_n)^{1/2} \right) \sum_{j=1}^{p_n} R_j^2 / R_0^2$ is large enough. Let $N_n$ be the number of correlation coefficients $R_j^2 / R_0^2 \geq \rho_n^2$ for $j \in [1, p_n]$, so that $\left( n / (2p_n)^{1/2} \right) \sum_{j=1}^{p_n} R_j^2 / R_0^2 \geq n N_n \rho_n^2 / (2p_n)^{1/2}$. The Box-Pierce test is consistent if

$$n^{1/2} \left( \frac{N_n}{p_n^{1/2}} \right)^{1/2} \rho_n \to \infty, \tag{1.3}$$

a condition which allows for $\rho_n = o\left(n^{-1/2}\right)$ provided there are enough correlation coefficients larger than $\rho_n$, that is, $N_n / p_n^{1/2} \to \infty$, which holds in particular when the exact order of $N_n$ is $p_n$. In other words, summing squared sample correlations in the Box-Pierce statistic allows us to detect very small population correlations provided they are not too sparse and are concentrated at lags smaller than $p_n$. As shown in this paper, such alternatives are not detected by Cramér-von Mises tests.

An important limitation of the critical region (1.1) is the use of an ad hoc order $p_n$. Many authors consider a deterministic $p_n$ such that $p_n \to \infty$. This choice of order is inadequate for detecting alternatives with correlations at low lags: taking $p_n = 30$ for instance is unlikely to give a test with power against popular $AR(1)$ or $MA(1)$ alternatives on samples of moderate size. Conversely, taking a fixed $p_n$ is not suitable for detecting higher order alternatives.

The need to properly address the tuning of a smoothing parameter with a role similar to $p_n$ has spurred the development of data-driven approaches for various nonparametric testing problems. The so-called adaptive approach, focuses on data-driven tests which detect alternatives in a smoothness class converging to the null at the fastest possible rate given that the smoothness class is unknown to the test user. See in particular Fan (1996), Spokoiny (1996), Horowitz and Spokoiny (2001), Guerre and Lavergne (2005), Guay and Guerre (2006) and Chen and Gao (2007) for various nonparametric models and related null hypotheses of theoretical or practical relevance. Golubev, Nussbaum and Zhou (2010) has proved Le Cam equivalence of Gaussian time series with spectral densities in a Besov space with the continuous-time Gaussian white noise model considered in Spokoiny (1996). This result is limited to Gaussian time series and is not useful in practice since it does not deliver ready-to-apply white noise tests. In fact, most of the data-driven choices of $p_n$ proposed in the white noise testing literature are not adaptive rate-optimal. As an exception, Fan and Yao (2005) extend the work of Fan (1996), outlining but not analyzing a data-driven test which is based on the maximum of a set of standardized Box-Pierce statistics, see also Golubev et al. (2010).

A popular data-driven method of choosing the order is the selection procedure proposed by Newey and West (1994) in the context of long run variance estimation. See, among other, the simulation section of Hong and Lee (2005). This selection procedure is however difficult to justify theoretically. Newey and West selection method, although being optimal for long-run variance estimation, does not produce a rate-optimal test because the optimal order for testing differs from the optimal order for estimation, see e.g. Guerre and Lavergne (2002) and the references therein. Escanciano and Lobato (2009) study a data-driven choice of order based on an AIC/BIC criterion which is suitable for estimation but is not adaptive rate-optimal for tests of the white noise hypothesis. This contrasts with the new data-driven tests proposed here.

The paper is organized as follows. Section 2 describes the penalty approach leading to the data-driven order $\widehat{p}$ and the construction of the rejection region of the test. Section 3 studies the statistical properties of the test under the general weak white noise null hypothesis and under the new class of alternatives mentioned above. It illustrates the importance of the choice of a suitable penalty both under the null and the alternative. Section 4 states our adaptive rate-optimality results and compares the new test with the Cramér-von Mises test of Deo (2000), the data-driven test of Escanciano and Lobato (2009) and the maximum test of Xiao and Wu (2011). Section 5 reports a simulation experiment that proposes a calibration of the penalty term and compares our automatic test with other data-driven tests, including tests of Deo (2000) or Escanciano and Lobato (2009) and a test that uses the Newey and West (1994) plug-in order selection procedure. Section 6 concludes. Proofs can be found in the supplementary material.

## 2. Construction of the test and choice of the critical values

Consider a variable $u_t$, $t = 1, ..., n$, which is either directly observed or defined as the error of a parametric model $m(X_t; \theta) = u_t$ with some observed covariate $X_t$. In the later case $u_t$ is not observed but can be estimated using the residuals $\widehat{u}_t = u_t(\widehat{\theta})$ where $\widehat{\theta}$ is an estimator of $\theta$. We are interested in testing that $u_t$ is uncorrelated. Suppose $\{u_t\}$ is a stationary process with zero mean and covariance function $R_j = \mathrm{Cov}(u_t, u_{t+j})$. The null and alternative hypotheses are then

$$\mathcal{H}_0 : R_j = 0 \text{ for all } j \neq 0 \qquad \text{versus} \qquad \mathcal{H}_1 : R_j \neq 0 \text{ for some } j \neq 0.$$

A natural estimator of the covariance is $\widehat{R}_j = \sum_{t=1}^{n-|j|} \widehat{u}_t \widehat{u}_{t+|j|}/n$, $j = 0, \pm 1, \ldots, \pm(n-1)$, which uses the residuals as if they were the true error terms. Given the kernel spectral density estimator

$$\hat{f}_n(\lambda; p) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} K\left(\frac{|j|}{p}\right) \widehat{R}_j \exp\left(-ij\lambda\right), \qquad K(0) = 1 \qquad \text{and} \int_0^{\infty} K(x)\, dx = 1,$$

where the support of $K$ is $[0,1]$, Hong (1996) has proposed a test statistic

$$\widehat{S}_p = n\pi \int_{-\pi}^{\pi} \left| \hat{f}_n(\lambda;p) - \frac{\widehat{R}_0}{2\pi} \right|^2 d\lambda = n \sum_{j=1}^{n-1} K^2\left(\frac{j}{p}\right) \widehat{R}_j^2. \tag{2.1}$$

For the uniform kernel $K(t) = \mathbb{I}(t \in [0,1])$ and up to a division by $\widehat{R}_0^2$, $\widehat{S}_p$ is the Box-Pierce statistic $\widehat{BP}_p/\widehat{R}_0^2 = n\sum_{j=1}^{p} \widehat{R}_j^2/\widehat{R}_0^2$. Large values of $\widehat{S}_p$ indicate evidence against the null. Under certain weak dependence conditions on the weak white noise $\{u_t\}$ and for $p = p_n \to \infty$ growing with a suitable rate, Shao (2011b) shows that $\left(\left(\widehat{S}_p - \widehat{S}_1\right)/R_0^2 - E_\Delta(p)\right)/V_\Delta(p)$ converges to a standard normal where

$$E_\Delta(p) = \sum_{j=1}^{n-1}\left(1 - \frac{j}{n}\right)\left(K^2\left(\frac{j}{p}\right) - K^2(j)\right),$$

$$V_\Delta^2(p) = 2\sum_{j=1}^{n-1}\left(1 - \frac{j}{n}\right)^2\left(K^2\left(\frac{j}{p}\right) - K^2(j)\right)^2,$$

and we shall use accordingly $E_\Delta(p)$ and $V_\Delta^2(p)$ as a standardization for $\left(\widehat{S}_p - \widehat{S}_1\right)/R_0^2$. In this notation, the subscript "$\Delta$" indicates difference $\widehat{S}_p - \widehat{S}_1$. For the Box-Pierce statistic, $E_\Delta(p) = (p-1)(1 + O(p/n))$ and $V_\Delta^2(p) = 2(p-1)(1+O(p/n))$ and these approximations remain valid for other kernels up to a multiplicative constant. We propose to select $\widehat{p}$ as the smallest integer number maximizing the penalized statistic,

$$\begin{aligned}\widehat{p} &= \arg\max_{p\in[1,\overline{p}_n]} \left(\frac{\widehat{S}_p}{\widehat{R}_0^2} - E(p) - \gamma_n V_\Delta(p)\right) \\ &= \arg\max_{p\in[1,\overline{p}_n]} \left(\frac{\widehat{S}_p - \widehat{S}_1}{\widehat{R}_0^2} - E_\Delta(p) - \gamma_n V_\Delta(p)\right),\end{aligned} \tag{2.2}$$

where $E(p) = \sum_{j=1}^{n-1}(1 - j/n) K^2(j/p)$ and $\overline{p}_n \leq n-1$. This penalization procedure is similar to penalization proposed by Guay and Guerre (2006) or Guerre and Lavergne (2005). It differs from the penalization used in the AIC or BIC procedures which use a higher penalty term $\gamma_n E(p)$ in place of $E(p) + \gamma_n V_\Delta(p)$. Escanciano and Lobato (2009) similarly use penalty term $\widehat{\gamma}_n E(p)$ for $p$ in a bounded finite set.

The intuition for $\widehat{p}$ is as follows. Note first that (2.2) uses the difference $\widehat{S}_p - \widehat{S}_1$. The idea here is that the test should be based on $\widehat{S}_1$ unless $\widehat{S}_p - \widehat{S}_1$ is large enough for some $p$. Since the criterion maximized in (2.2) is equal to 0 for $p = 1$, $\widehat{p}$ differs from 1 whenever there is a $p$ such that $\left(\widehat{S}_p - \widehat{S}_1\right)/\widehat{R}_0^2 - E_\Delta(p) - \gamma_n V_\Delta(p) > 0$ or equivalently

$$\frac{\left(\widehat{S}_p - \widehat{S}_1\right)/\widehat{R}_0^2 - E_\Delta(p)}{V_\Delta(p)} > \gamma_n, \tag{2.3}$$

an inequality which, in view of the asymptotic normality established by Shao (2011b) under the null, has the flavour of a one-sided significance test using a critical value $\gamma_n$. Such a construction suggests that the data-driven statistic $\widehat{S}_{\widehat{p}}$ better captures higher order covariances than $\widehat{S}_1$. Therefore, rejecting the null when $\widehat{S}_{\widehat{p}} \geq z$ should give a more powerful test than the test $\widehat{S}_1 \geq z$ based on $\widehat{S}_1$ and the same critical value $z$ as recommended below. See (3.8) in Theorem 4 for a more formal statement. Why the chosen $\widehat{p}$ should have certain optimality properties can be seen by viewing (2.2) as a bias variance trade-off. Theorem 2.2 in Shao (2011b) suggests that $\left(\widehat{S}_p - \widehat{S}_1\right)/\widehat{R}_0^2 - E_\Delta(p)$ is an estimator of $n\sum_{j=2}^{\infty} R_j^2$ with a bias $n\sum_{j=p+1}^{\infty} R_j^2$ and a standard deviation $V_\Delta(p)$. Hence (2.2) choses a $p$ which maximizes $-n\sum_{j=p+1}^{\infty} R_j^2 - \gamma_n V_\Delta(p)$ and therefore achieves the so called bias variance trade-off, leading to a data-driven test statistic $\widehat{S}_{\widehat{p}} = \widehat{S}_1 + \widehat{S}_{\widehat{p}} - \widehat{S}_1$ with the best potential to detect an alternative.

Under $\mathcal{H}_0$, it is expected that $\widehat{p} = 1$ with a high probability provided $\gamma_n$ is large enough since all the $\widehat{S}_p - \widehat{S}_1$ estimate 0. Since $\widehat{S}_{\widehat{p}} = \widehat{S}_1 + o_{\mathbb{P}}(1)$ under the null, the critical values of the test can be taken to be the same as the critical values of the test based upon the simple statistic $\widehat{S}_1$. A HAC-robust standardization of $\widehat{S}_1$ is given in Lee (2007). In the case where $u_t$ is observed, an inconsistent "estimator" of the long run variance of $\sum_{t=1}^{n-1} u_t u_{t+1}/(n-1)$

is, for a kernel $k(\cdot)$, $k_{ij} = k(|i-j|/n)$ and $\varphi_i = \sum_{t=1}^{i-1}\left(u_t u_{t+1} - \widehat{R}_1\right)/n^{1/2}$,

$$\widetilde{\Gamma}_1 = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left(\left(k_{ij} - k_{i,j+1}\right) - \left(k_{i+1,j} - k_{i+1,j+1}\right)\right)\varphi_i\varphi_j.$$

For residuals $\hat{u}_t$, let $\widehat{\theta}_i$ be the estimator $\widehat{\theta}$ computed with the first $i$ observations and estimate $\varphi_i$ recursively by $\widehat{\varphi}_i = \sum_{t=1}^{i-1}\left(u_t\left(\widehat{\theta}_i\right)u_{t+1}\left(\widehat{\theta}_i\right) - \widehat{R}_1\right)/n^{1/2}$. Let

$$\widehat{\Gamma}_1 = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left(\left(k_{ij} - k_{i,j+1}\right) - \left(k_{i+1,j} - k_{i+1,j+1}\right)\right)\widehat{\varphi}_i\widehat{\varphi}_j.$$

It follows from Lee (2007) that the limit distribution of $n\widehat{R}_1/\widetilde{\Gamma}_1$ when $u_t$ is observed and of $n\widehat{R}_1/\widehat{\Gamma}_1$ when $u_t$ is is estimated by residuals $\hat{u}_t$ is, assuming that $k(\cdot)$ is twice continuously differentiable

$$\frac{W^2(1)}{-\int_0^1\int_0^1 k''(r-s)\left(W(r) - rW(1)\right)\left(W(s) - sW(1)\right)drds} \tag{2.4}$$

where $W$ is a standard Brownian motion. Let $z_L(\alpha)$ be be the $(1-\alpha)$th quantile of (2.4). The critical values and rejection region of the test are

$$\widehat{z}_L(\alpha) = K^2(1)\,\widetilde{\Gamma}_1 z_L(\alpha), \tag{2.5}$$

$$\widehat{z}_{KL}(\alpha) = K^2(1)\,\widehat{\Gamma}_1 z_L(\alpha), \tag{2.6}$$

$$\widehat{S}_{\widehat{p}} \geq \widehat{z}(\alpha) \quad \text{where } \widehat{z}(\alpha) = \begin{cases} \widehat{z}_L(\alpha) & \text{for observed } \{u_t\}, \\ \widehat{z}_{KL}(\alpha) & \text{for residuals } \{\widehat{u}_t\}. \end{cases} \tag{2.7}$$

We also consider a modified version of the test which employs a standardization of the sample covariances as used by Deo (2000) or Escanciano and Lobato (2009),

$$\widehat{S}_p^* = n\sum_{j=1}^{n-1}K^2\left(\frac{j}{p}\right)\left(\frac{\widehat{R}_j}{\widehat{\tau}_j}\right)^2 \quad \text{where } \widehat{\tau}_j^2 = \frac{1}{n-j}\sum_{t=1}^{n-j}\widehat{u}_t^2\widehat{u}_{t+j}^2 - \left(\frac{n}{n-j}\widehat{R}_j\right)^2. \tag{2.8}$$

The sample variance $\widehat{\tau}_j^2$ is an estimator of $\tau_j^2 = \operatorname{Var}(u_t u_{t+j})$ which, for observed $u_t$, is the asymptotic variance of $n^{1/2}\left(\widehat{R}_j - R_j\right)$ in the case of uncorrelated $u_t u_{t+j}$ or for martingale difference. The corresponding data-driven order $p$ and critical values are

$$\widehat{p}^* = \arg\max_{p\in[1,\overline{p}_n]} \left(\widehat{S}_p^* - E(p) - \gamma_n V_\Delta(p)\right), \tag{2.9}$$

$$\widehat{z}^*(\alpha) = \frac{\widehat{z}(\alpha)}{\widehat{\tau}_1^2}. \tag{2.10}$$

While the test (2.7) is studied in Theorems 1 and 2, the test with rejection region $\widehat{S}_{\widehat{p}^*}^* \geq \widehat{z}^*(\alpha)$ is studied in Theorem 3.

Let us now turn to notations and our main assumptions. In what follows, $a_n \asymp b_n$ means that the sequences $\{a_n\}$ and $\{b_n\}$ have the same order, i.e. that $a_n/b_n$ and $b_n/a_n$ are both $O(1)$. For a real random variable $Z$ and a positive real number $a$, $\|Z\|_a = \mathbb{E}^{1/a}[|Z|^a]$. Consider first the case of observed $u_t$. When studying the performance of the test under the alternative, we consider a sequence $\{u_{t,n}\}$ of stationary alternatives with autocovariance coefficients $\{R_{j,n}\}$. This means that for each given $n$, the process $\{u_{t,n}, t \in \mathbb{N}\}$ is stationary. This type of sequences includes for instance local $MA(\infty)$ alternatives $u_{t,n} = \varepsilon_t + \sum_{i=1}^{\infty} a_{i,n}\varepsilon_{t-i}$ where $a_{i,n} \to 0$ when $n$ grows. Further, for residuals $\widehat{u}_t = u_t\left(\widehat{\theta}\right)$, we assume that $\sqrt{n}\left(\widehat{\theta} - \theta_n\right)$ is asymptotically centered with $\theta_n$ is a pseudo-true value and set $u_t(\theta_n) = u_{t,n}$. For the sake of brevity, $\{u_{t,n}\}$ and $\{R_{j,n}\}$ are abbreviated to $\{u_t\}$ and $\{R_j\}$ in the rest of the paper but we maintain the dependence with respect to $n$ when stating our main assumptions. Under the null and the alternative, we follow Shao (2011b), Xiao and Wu (2011), and restrict ourselves to stationary processes satisfying a moment contraction condition by Wu (2005). We assume that $u_{t,n} = F_n(\ldots, e_{t-1}, e_t)$ for some measurable $F$, where $e_t$, $t = -\infty, \ldots, +\infty$, are i.i.d. (univariate or vector) random variables. Consider an independent copy $\{e'_t\}$ of $\{e_t\}$ and define

$$u_{t,n}^{\tau} = F_n(\ldots, e_{\tau-1}, e'_{\tau}, e_{\tau+1}, \ldots, e_{t-1}, e_t) \qquad \tau \leq t \leq n,$$

where $e_\tau$ is changed to $e'_\tau$. Assume that for some $a > 0$ and for all $j \geq 0$,

$$\left\|u_{t,n} - u_{t,n}^{t-j}\right\|_a \leq \delta_a(j) \qquad \text{where } \delta_a(j) \to 0 \text{ when } j \to \infty,$$

a condition meaning that shocks cannot have a long run impact. A fast decrease of $\delta_a(j)$ also ensures that $u_t = u_{t,n}$ becomes independent of $u_{t-j}$ when $j$ grows as the $\alpha$-mixing assumption used in Francq et al. (2005) or Delgado and Velasco (2012). Shao (2011b) assumes that $\delta_a(j)$ decreases at an exponential rate, a condition which is satisfied by many linear and nonlinear time series models, including threshold, stochastic volatility, bilinear or GARCH models, see Shao (2011b), Wu (2005, 2007) and the references therein. Our main assumptions are given below.

**Assumption K** (Kernel). *The kernel function $K(\cdot)$ in (2.1) from $\mathbb{R}^+$ to $[0,\infty)$ is non-increasing, bounded away from 0 on $[0,1/2]$ and continuous differentiable over its support $[0,1]$. The kernel $k(\cdot)$ used for the critical values is twice continuously differentiable over its compact support.*

**Assumption R** (Regularity). *Under $\mathcal{H}_0$ and $\mathcal{H}_1$, $\sup_t \|u_{t,n}\|_{12a} < C_0 R_{0,n}^{1/2}$ for some $a > 1$ and, for some $b > 0$, $\delta_{12a}(j) \leq C_1 j^{-7-b}$. Moreover $1/C_2 \leq R_{0,n} \leq C_2$, and $\max_{j\in[1,\overline{p}_n]} R_{0,n}^2 / \operatorname{Var}(u_{t,n}u_{t+j,n}) \leq C_3$.*

**Assumption P** (Order $p$). *The maximal order $\overline{p}_n$ diverges faster than some power of $n$ with $\overline{p}_n = o(n^{1/(2(1+3/a))})$ as $n \to \infty$, where $a > 1$ is the same constant as in Assumption R above. The penalty sequence $\gamma_n$ satisfies $\gamma_n > 0$, $\gamma_n \to \infty$ and $\gamma_n = o\left(n^{1/4}\right)$ as $n \to \infty$.*

**Assumption M** (Model). *The processes $\{u_{t,n}\}$, the model $m(X_t;\theta) = u_t$ and the estimators $\left\{\widehat{\theta}_t\right\}$ satisfy the following conditions:*

*(i) There is a sequence $\{\theta_n\}$, with $\theta_n = \theta_0$ for all $n$ under $\mathcal{H}_0$, such that*

$$\left\{ \left( n^{1/2}\left(\widehat{\theta}_{[ns]} - \theta_n\right)', n^{-1/2}\sum_{t=1}^{[ns]}\left(u_{t,n}u_{t-1,n} - \mathbb{E}\left[u_{t,n}u_{t-1,n}\right]\right)\right)', s \in [0,1] \right\} \tag{2.11}$$

*$D_{[0,1]}$-converges in distribution to a Brownian motion with a full rank covariance matrix.*

*(ii) The residual function admits a second order expansion* $u_t(\theta) = u_{t,n} + (\theta - \theta_n)' u_{t,n}^{(1)} + (\theta - \theta_n)' u_{t,n}^{(2)} (\theta - \theta_n) + \mathfrak{r}_{t,n}(\theta)$ *where, for any* $C > 0$,

$$\sup_{t \in [1,n]} \sup_{\theta; \|\theta - \theta_n\| \le C n^{-1/2}} |\mathfrak{r}_{t,n}(\theta)| = o_{\mathbb{P}}\left(\frac{1}{n}\right) \tag{2.12}$$

*and, for each* $n$, $\{u_{t,n}, u_{t,n}^{(1)}, u_{t,n}^{(2)}\}$ *is a stationary process with* $\mathbb{E}^{1/2}\left[\|a_t\|^2\right] \le C_4$, $\{a_t\}$ *being successively* $\left\{u_{t,n}^{(1)}\right\}$, $\left\{u_{t,n}^{(2)}\right\}$ $\{u_{t,n}^2\}$, $\left\{u_{t,n} u_{t,n}^{(1)}\right\}$, $\left\{u_{t,n}^{(1)} u_{t,n}^{(1)\prime}\right\}$, $\left\{u_{t,n} u_{t,n}^{(2)}\right\}$, *and where* $\sum_{j=-\infty}^{\infty} \mathbb{E}\left[\left\|u_{t-j,n}^{(1)} u_{t,n}\right\|^2\right] \le C_5$, $\sup_{j \in \mathbb{Z}} \mathbb{E}\left[\left\|n^{-1/2} \sum_{t=j+1}^{n} \left(u_{t-j,n}^{(1)} u_{t,n} - \mathbb{E}[u_{t-j,n}^{(1)} u_{t,n}]\right)\right\|^2\right] \le C_6$, $\sup_{j \in \mathbb{Z}} \mathbb{E}\left[\left\|u_{t,n}^{(1)} u_{t,n} u_{t-j,n}^2\right\|\right] \le C_7$, *and* $\sup_{j \in \mathbb{Z}} \mathbb{E}\left[\left\|n^{-1/2} \sum_{t=j+1}^{n} \left(u_{t,n}^{(1)} u_{t,n} u_{t-j,n}^2 - \mathbb{E}[u_{t,n}^{(1)} u_{t,n} u_{t-j,n}^2]\right)\right\|^2\right] \le C_8$.

The compact sets $[0, 1/2]$ and $[0, 1]$ in Assumption K are somehow arbitrary and can be replaced by any nested compact intervals. Note however that Assumption K forbids the use of the Daniell kernel $K(x) = \sin(x)/x$ due to the nonincreasing function and bounded support conditions.

Assumption R imposes a polynomial decay on the coefficients $\delta_{12a}(j)$, a condition which is weaker than the exponential rate assumed in Shao (2011b). Note that in Assumption P the order of $\overline{p}_n$ can come closer to $n^{1/2}$ when $a$ is high, that is when $u_t$ has finite moments of higher order. Under Assumption R, $\{u_{t,n}\}$ must have finite moments of order twelve at least. This is mostly needed for a proof of Theorem 1 below based on Lindeberg substitution method, see Pollard (2002, p. 179), which uses moment bounds as the Cauchy-Schwarz inequality $\mathbb{E}\left[\left(u_t^2 u_{t+j}^2\right)^3\right] \le \mathbb{E}\left[u_t^{12}\right]$. Since implementing the proposed data-driven tests with a large $\overline{p}_n$ would in principle allow us to detect a wider class of alternatives, Assumption P, which plays an important role under the null in our proofs, may be too restrictive. Our simulation experiments indeed suggest that Assumption P can be weakened when focusing on white noise processes of practical relevance since the order $\overline{p}_n \asymp n$ gives good results for various white noise processes of practical interest. On the other hand, choosing a smaller $\overline{p}_n$

still gives a good power, see comments on Table 5 at the end of the simulation experiments section.

When $\{u_t\}$ is observed, Assumption M is equivalent to Assumption 1 of Lobato (2001) and the FCLT for $n^{-1/2}\sum_{t=1}^{[ns]}(u_t u_{t-1} - \mathbb{E}[u_t u_{t-1}])$ is a consequence of Assumption R and the FCLT of Wu (2007). Assumption M is easily verified for simple linear models and OLS estimation where $u_{t,n}^{(2)}$ and $\mathfrak{r}_{t,n}$ can be set to 0. Assumption M-(i) is a shortened version of Assumptions B1 and A2 of Kuan and Lee (2006) who employ a standard linear expansion $n^{1/2}\left(\widehat{\theta} - \theta_n\right) = n^{-1/2}\sum_{t=1}^{n}\ell_t + o_{\mathbb{P}}(1)$ to show that (2.11) satisfies a functional central limit theorem (FCLT) called for in M-(i). The FCLT is mostly used under $\mathcal{H}_0$ to show that $\mathbb{P}\left(\widehat{S}_1 \geq \widehat{z}(\alpha)\right) \to \alpha$ and $\mathbb{P}\left(\widehat{S}_1^* \geq \widehat{z}^*(\alpha)\right) \to \alpha$ in the case of residuals. The full-rank FCLT condition in Assumption M-(i) implies certain restrictions. For example, for a correctly specified $AR(1)$ model $X_t - \theta X_{t-1} = u_t$, the case of $\theta = 0$ is ruled out, a value of the parameter which would in principle be excluded when considering such an $AR(1)$ specification. Theorem 4 at the end of the next section explains how to overcome this issue with an alternative choice of critical values when Assumption M-(i) is too restrictive. The next section describes some suitable theoretical requirements for the penalty sequence $\gamma_n$ while the simulation section proposes a calibration of $\gamma_n$ which gives good results for various white noise processes and alternatives.

## 3. Asymptotic level and consistency

An important issue in the construction of the test (2.7) is the choice of the penalty sequence. Choosing $\gamma_n$ large enough implies that $\widehat{p}$ stays close to 1 and so the test statistic $\widehat{S}_{\widehat{p}}$ remains close to $\widehat{S}_1$. Hence, on the one hand, using large $\gamma_n$ ensures that the level of the test is close to its nominal size. On the other hand, a large $\gamma_n$ may substantially limit the power of the test since the statistic $\widehat{S}_{\widehat{p}}$ would not differ from $\widehat{S}_1$. The trade-off between size and power is addressed by Theorem 1 and Theorem 2.

Consider first the properties of the test under the null hypothesis. The following theorem gives a lower bound for $\gamma_n$ which ensures that $\widehat{p} = 1$ asymptotically so that the test is asymptotically of level $\alpha$.

**Theorem 1.** *Let Assumptions K , M, P and R hold. If the penalty sequence* $\{\gamma_n, n \geq 1\}$ *satisfies*

$$\gamma_n \geq (1+\epsilon)(2\ln\ln n)^{1/2} \quad \text{for some } \epsilon > 0, \tag{3.1}$$

*then under* $\mathcal{H}_0$, $\lim_{n\to\infty} \mathbb{P}(\widehat{p} = 1) = 1$ *and the test (2.7) is asymptotically of level* $\alpha$.

Under the null hypothesis, the selected order $\widehat{p}$ is asymptotically equal to 1. It follows that $\widehat{S}_{\widehat{p}} = \widehat{S}_1 + o_{\mathbb{P}}(1)$ and that critical values (2.5) or (2.6) guarantee that the test is asymptotically of level $\alpha$. A key result is therefore that $\lim_{n\to\infty} \mathbb{P}(\widehat{p} = 1) = 1$ holds under various white noise models and observed $u_t$ or residuals $\widehat{u}_t$. That the estimation has no impact asymptotically follows from (3.1) which imposes $\gamma_n \to \infty$. When $\widehat{\theta}$ is $\sqrt{n}$-consistent, estimating the residuals gives test statistics satisfying

$$\widehat{S}_p = n \sum_{j=1}^{p} \left( \frac{1}{n} \sum_{t=1}^{n-j} u_t u_{t+j} \right)^2 + O_{\mathbb{P}}(1)$$

uniformly in $p$. The fact that the remainder term $O_{\mathbb{P}}(1)$ is negligible compared to $\gamma_n$ is a crucial element in showing that the asymptotic behavior of $\widehat{p}$ is not affected by the estimation under the null. The divergence of $\gamma_n$ is also important to account for the fact that the standardization $E_{\Delta}(p)$ and $V_{\Delta}(p)$ are only valid when $p \to \infty$ since $\gamma_n \to \infty$ imposes that either $\widehat{p} = 1$ or $\widehat{p}$ diverges because (2.3) cannot hold for finite $p > 1$. Compared to the existing adaptive results of Horowitz and Spokoiny (2001), Guerre and Lavergne (2005), Guay and Guerre (2006) or Chen and Gao (2007), an important technical contribution of our paper is that Theorem 1 holds without assuming that the set of admissible $p$ is a power set $\{a^j, j \in \mathbb{N}\}$, $a > 1$.

Another important finding is that the penalty sequence $\gamma_n$ can diverge with the low order $(\ln\ln n)^{1/2}$ allowed by (3.1). This contrasts with the larger order $\ln n$ used in the BIC selection procedure and in the corresponding data-driven tests. In view of the potential negative impact of a large $\gamma_n$ on the power of the test, it is worth asking if the lower bound (3.1) can be improved, that is if $\mathbb{P}(\widehat{p}=1)\to 1$ would be ensured for even lower values of penalty term $\gamma_n$. The proof suggests that this is not the case. The main argument is based on expression

$$\mathbb{P}(\widehat{p}\neq 1)=\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\left(\frac{\left(\hat{S}_p-\hat{S}_1\right)/\hat{R}_0^2-E_\Delta(p)}{V_\Delta(p)}\right)\geq\gamma_n\right) \tag{3.2}$$

for the probability of not selecting 1. It can be seen from the proof of Theorem 1 that, for the Box-Pierce version of the test, the right-hand side of (3.2) asymptotically behaves like the maximum of standardized partial sums whose exact order is $(2\ln\ln n)^{1/2}$, see (B.38) in the Supplementary Material. Hence the bound (3.1) is optimal to achieve $\mathbb{P}(\widehat{p}=1)\to 1$.

Let us now turn to the detection properties of the test. Recall that the covariance of the alternative may depend on the sample size so that $R_j=R_{j,n}$ may go to 0 when $n$ increases. The new class of alternatives is defined similarly to (1.3) in the introduction section. Consider first a sequence $\rho_n\to 0$ and a lag order $P_n$. An important indicator for detection of alternatives is the number of correlations above $\rho_n$,

$$N_n=N_n(P_n,\rho_n)=\#\left\{|R_j/R_0|\geq\rho_n,\quad 1\leq j\leq P_n\right\}. \tag{3.3}$$

The next theorem gives a detection condition on $N_n$, $P_n$ and $\rho_n$.

**Theorem 2.** *Suppose Assumptions K, M, R and P hold. There exists a constant $\kappa_*>0$ such that the test (2.7) is consistent against all alternatives $\{u_t\}$ satisfying, for some $\rho_n>0$ and $P_n\in[1,\overline{p}_n/2]$,*

$$n^{1/2}\left(\frac{N_n}{\gamma_n P_n^{1/2}}\right)^{1/2}\rho_n\geq\kappa_*. \tag{3.4}$$

Condition (3.4) is similar to the detection condition (1.3) required for consistency of the Box-Pierce test (1.1). However a key difference between the two conditions is that while in (1.3) the lag order $p_n$ is assumed known and is used in the construction of the test statistic, in (3.4) the lag order $P_n$ in (3.4) is unknown. This illustrates the adaptive capability of the new test. A second important difference between (1.3) and (3.4) is that the latter involves penalty sequence $\gamma_n$. For given $P_n$ and $N_n$ detection condition (3.4) admits rate $\rho_n^*$ satisfying

$$\rho_n^* \asymp \frac{1}{n^{1/2}} \left( \frac{\gamma_n P_n^{1/2}}{N_n} \right)^{1/2} . \tag{3.5}$$

Rate $\rho_n^*$ in (3.5) deteriorates with the penalty sequence. Condition (3.4) thus demonstrates the potential negative impact of the penalty sequence on the power of the test. This impact can also be seen from proof of Theorem 2 which uses the fact that the test (2.7) rejects the null whenever

$$\frac{\widehat{S}_p - \widehat{R}_0^2 E(p)}{\widehat{R}_0^2 V_\Delta(p)} \geq \gamma_n + \frac{\widehat{z}(\alpha)}{\widehat{R}_0^2 V_\Delta(p)} \text{ for some } p \in [2, \overline{p}_n] . \tag{3.6}$$

For the alternatives for which (3.6) only holds for $p \to \infty$ so that $V_\Delta(p) \to \infty$, (3.6) suggests that $\gamma_n$ may be more important than the critical value $\widehat{z}(\alpha)$ regarding detection.

Two special cases of (3.5) are worth mentioning. First, the situation where $\lim_{n\to\infty} \gamma_n P_n^{1/2}/N_n = 0$ is of special interest since (3.5) shows that the test can detect correlation coefficients converging to 0 at a rate that is faster than the parametric rate $n^{-1/2}$. The best possible rate in this case is $\rho_n^* \asymp \gamma_n^{1/2} / \left( n P_n^{1/2} \right)^{1/2}$ which is achieved for "saturated" alternatives with $N_n \asymp P_n$. Second, a less favorable case corresponds to more sparse correlation coefficients satisfying $\lim_{n\to\infty} \gamma_n P_n^{1/2}/N_n = \infty$. In this case (3.5) does not allow for correlation coefficients converging to 0 at the rate of $n^{-1/2}$. This case has been covered by Donoho and Jin (2004) for a theoretical model where a known number $P_n$ of independent Gaussian variables with mean $n (R_j/R_0)^2$ and variance 1 is observed. These authors show that in such a setup the best possible detection rate is $\rho_n = (\ln n/n)^{1/2}$, a rate which is achieved by the maximum white noise test of Xiao and Wu (2011). This suggests that our test may not be optimal when $\lim_{n\to\infty} \gamma_n P_n^{1/2}/N_n = \infty$. However, it is shown in Proposition 1 in Section 4 below that the test of Xiao and Wu (2011), unlike our test, does not detect moderately sparse alternatives satisfying (3.5) with $\lim_{n\to\infty} \gamma_n P_n^{1/2}/N_n = 0$ and $\gamma_n \asymp (2 \ln \ln n)^{1/2}$.

We conclude this section with two extensions of our main results. The first extension shows that the test derived from (2.8) and (2.9) has similar properties as the test (2.7).

**Theorem 3.** *Suppose Assumptions K, M, R and P hold. Then $\mathbb{P}(\widehat{p}^* = 1) \to 1$ under $\mathcal{H}_0$ and the test which rejects the null when $\widehat{S}^*_{\widehat{p}^*} \geq \widehat{z}^*(\alpha)$ is asymptotically of level $\alpha$. It also detects the alternatives satisfying (3.4) in Theorem 2 for a large enough $\kappa_*$.*

The second extension is useful in the case of residuals when the full-rank FCLT condition in Assumption M-(i) is too restrictive so that the critical value $\widehat{z}_{KL}(\alpha)$ in (2.6) cannot be used. Suppose that an additional test statistic $\widehat{T}_n$ with critical values $\widehat{t}_n(\alpha)$ satisfying $\lim_{n\to\infty} \mathbb{P}\left(\widehat{T}_n \geq \widehat{t}_n(\alpha)\right) = \alpha$ under the null is available. Consider the critical value

$$\widehat{c}^*_n(\alpha) = \widehat{S}^*_1 - \widehat{T}_n + \widehat{t}_n(\alpha). \tag{3.7}$$

**Theorem 4.** *Suppose that Assumptions K, R and P hold, as Assumption M-(ii) with $\sqrt{n}\left(\widehat{\theta} - \theta_n\right) = O_{\mathbb{P}}(1)$ where the deterministic sequence $\{\theta_n\}$ is such that $\theta_n = \theta_0$ for all $n$ under $H_0$. Suppose also that (A0) $\lim_{n\to\infty} \mathbb{P}\left(\widehat{T}_n \geq \widehat{t}_n(\alpha)\right) = \alpha$ under $\mathcal{H}_0$ and (A1) $\widehat{c}_n(\alpha) \leq O_{\mathbb{P}}(\gamma_n)$ under the considered alternative. Then the test which rejects the null when $\widehat{S}^*_{\widehat{p}^*} \geq \widehat{c}_n(\alpha)$ is asymptotically of level $\alpha$ and detects the alternatives satisfying the condition (3.4) of Theorem 2 for a sufficiently large $\kappa_*$. Moreover and even if (A1) does not hold, we have under the alternative and for any sample size $n$,*

$$\mathbb{P}\left(\widehat{S}^*_{\widehat{p}^*} \geq \widehat{c}_n(\alpha)\right) \geq \mathbb{P}\left(\widehat{T}_n \geq \widehat{t}_n(\alpha)\right). \tag{3.8}$$

Condition (A1), which allows for $\widehat{c}_n(\alpha) \xrightarrow{\mathbb{P}} -\infty$, means, when $\widehat{t}_n(\alpha) = O_{\mathbb{P}}(1)$ as usual, that $\widehat{T}_n$ diverges at least as fast as $\widehat{S}^*_1$ or that both lack power against the considered alternative and are $O_{\mathbb{P}}(1)$. The bound (3.8) means that the data-driven test is at least as powerful than the test based on $\widehat{T}_n$. As a consequence of (3.8), the test $\widehat{S}^*_{\widehat{p}^*} \geq \widehat{z}^*(\alpha)$ is as least as powerful as $\widehat{S}^*_1 \geq \widehat{z}^*(\alpha)$, $\widehat{z}^*(\alpha)$ as in (2.10). The use of the critical value (3.7) can give a data-driven test whose power properties can be tailored to be optimal against some specific alternatives by a proper choice of a corresponding optimal $\widehat{T}_n$. Examples of test statistic $\widehat{T}_n$ which does not require Assumption M-(i) can be found in Delgado and Velasco (2012) and Francq et al. (2005). Delgado and Velasco (2012) propose a Box-Pierce statistic corrected for estimation with an elegant general approach and some parametric optimality properties under Gaussianity whereas Francq et al. (2005) is more specific to ARMA specifications.

## 4. Adaptive rate-optimality and comparisons with other tests

While Theorem 1 gives the lower bound (3.1) of order $(2\ln\ln n)^{1/2}$ for the penalty sequence $\gamma_n$ that is necessary to ensure that the test is asymptotically of level $\alpha$, Theorem 2 suggests that increasing $\gamma_n$ can impair the power of the test. Hence a good compromise for the choice of the penalty sequence suitable both under $\mathcal{H}_0$ and $\mathcal{H}_1$ is $\gamma_n \asymp (2\ln\ln n)^{1/2}$. Once this choice is made one may ask if the resulting test is the best possible in the sense that there is no other test that can detect alternatives satisfying a condition less restrictive than (3.4), when $\kappa_* = \kappa_n \to 0$ is allowed. The absence of a better test is the so called adaptive rate-optimality. The next theorem establishes adaptive rate-optimality for alternatives satisfying $\lim_{n\to\infty} \gamma_n P_n^{1/2}/N_n = 0$.[1]

**Theorem 5.** *Let $u_t$ be observed. For any sequence $\kappa_n \to 0$, there exists a sequence of alternatives $\{u_t\}$ such that, for some $P_n \in [1, \overline{p}_n]$ and $\rho_n > 0$ with*

$$\rho_n \geq \frac{\kappa_n}{n^{1/2}} \left( \frac{(2\ln\ln n)^{1/2} P_n^{1/2}}{N_n} \right)^{1/2}, \qquad \lim_{n\to\infty} \frac{(2\ln\ln n)^{1/2} P_n^{1/2}}{N_n} = 0,$$

[1] As discussed when introducing approximation (3.5), the test (2.7) is not optimal for detection of sparse alternatives with $\lim_{n\to\infty} \gamma_n P_n^{1/2}/N_n = \infty$ which are not considered here.

*such that the other assumptions of Theorem 2 are satisfied, but that cannot be detected by any possible asymptotically $\alpha$-level test.*

Hence, when $\gamma_n \asymp (2\ln\ln n)^{1/2}$, it is not possible to improve on the detection condition (3.4) and the rate $\rho_n^*$ in (3.5) is optimal. We now give an example of alternatives which are detected by the test (2.7) but not by other popular tests. Consider the following high-order moving average process,

$$u_t = u_{t,n} = \varepsilon_t + \frac{\nu\gamma_n^{1/2}}{n^{1/2}P_n^{1/4}}\sum_{k=1}^{P_n}\psi_k\varepsilon_{t-k}, \qquad \sum_{k=1}^{P_n}\psi_k^2 = O(P_n), \qquad \lim_{n\to\infty} P_n = \infty, \tag{4.1}$$

where $\{\varepsilon_t\}$ is a strong white noise with variance $\sigma^2$, $\nu$ is a scaling constant and $\gamma_n \asymp (2\ln\ln n)^{1/2}$. This alternative has moving average coefficients of order $\gamma_n^{1/2}/\left(n^{1/2}P_n^{1/4}\right) = o\left(n^{-1/2}\right)$ provided $P_n$ diverges at a polynomial rate. Hence short term shocks have statistically negligible impact. However when $\psi_k = 1$ for all $k$, the long term multiplier of (4.1) is equal to $\nu\left(\gamma_n P_n^{3/2}/n\right)^{1/2}$ which is of larger order than $n^{-1/2}$. The following lemma describes the covariance function and conditional expectation of the alternative (4.1).

**Lemma 1.** *If $P_n = o((n/\gamma_n)^{2/3})$ and $\lim_{n\to\infty}(\gamma_n/n) = 0$ then the alternative $\{u_t\}$ in (4.1) satisfies $R_0 = \sigma^2\left(1 + O\left(\gamma_n P_n^{1/2}/n\right)\right)$ and, uniformly in $j \in [1, P_n]$,*

$$R_j = \frac{\nu\gamma_n^{1/2}}{n^{1/2}P_n^{1/4}}\psi_j\sigma^2 + o\left(\frac{\gamma_n^{1/2}}{n^{1/2}P_n^{1/4}}\right).$$

*Moreover*

$$\mathbb{E}\left[u_t|u_{t-k}, k \geq 1\right] = \frac{\nu\gamma_n^{1/2}}{n^{1/2}P_n^{1/4}}\sum_{k=1}^{P_n}\psi_k u_{t-k} + O_{\mathbb{P}}\left(\frac{\gamma_n P_n}{n}\right).$$

Hence a distinctive feature of the alternative (4.1) when $\max_{1\leq k\leq P_n}|\psi_k| = O(1)$ is that $\max_{j\geq 1}|R_j| = o\left(n^{-1/2}\right)$ provided $P_n/\gamma_n^2 \to \infty$. The expression of $\mathbb{E}\left[u_t|u_{t-k}, k \geq 1\right]$ reveals that $u_t$ can be very difficult to forecast since the coefficients of the lagged variables are all $o\left(n^{-1/2}\right)$ provided $P_n = o\left(n^{1/2}/\gamma_n\right)$. This suggests that such a process will be seen in practice as a martingale difference when using standard statistical tools. This may be a relevant example of alternatives in economical or financial contexts where arbitrage occurs.

We show in Proposition 1 below that the new tests detect these alternatives but that this is not the case for three tests based on the following test statistics,

$$W_n = b_n \left( n^{1/2} \max_{j \in [1,J_n]} \left| \frac{\widehat{R}_j}{\widehat{\tau}_j} \right| - b_n \right), \quad \text{where } b_n = \left( 2 \ln J_n - \ln \ln J_n - \ln (4\pi) \right)^{1/2}, \tag{4.2}$$

$$CvM_n = \frac{n}{\pi^2} \sum_{j=1}^{J_n} \frac{\widehat{R}_j^2}{j^2 \widehat{\tau}_j^2}, \tag{4.3}$$

$$EL_n = \widehat{BP}^*_{\widehat{p}^*_{EL}}, \quad \widehat{p}^*_{EL} = \arg \max_{p \in [1,J_n]} \left\{ \widehat{BP}^*_p - \widehat{\gamma}^*_{EL} p \right\} \text{ where} \tag{4.4}$$

$$\widehat{\gamma}^*_{EL} = \begin{cases} \ln n & \text{if } n^{1/2} \max_{j \in [1,J_n]} \left| \frac{\widehat{R}_j}{\widehat{\tau}_j} \right| \leq (2.4 \ln n)^{1/2}, \\ 2 & \text{otherwise.} \end{cases}$$

Statistic $W_n$ in (4.2) is studied in Xiao and Wu (2011) who show that $W_n$ asymptotically has an extreme value distribution. The statistic $CvM_n$ in (4.3), due to Deo (2000) for observed $u_t$, is a version of the Cramér-von Mises test of Durlauf (1991) partially corrected for heteroskedasticity. Test statistic $EL_n$ has been introduced by Escanciano and Lobato (2009) for observed $u_t$ and a fixed $J_n$. As in our test, the order $\widehat{p}^*_{EL}$ selected by Escanciano and Lobato (2009) is asymptotically equal to 1 under $\mathcal{H}_0$ and similar critical values can be used. To show that tests (4.2)–(4.4) do not detect alternatives with small correlation coefficients, it is sufficient to consider a Gaussian null hypothesis $G_0$ under which $\{u_t\}$ is a Gaussian white noise process $\{\varepsilon_t\}$ with variance $\sigma^2$ against an alternative $G_1$ under which $\{u_t\}$ is given by (4.1) with Gaussian i.i.d. $\{\varepsilon_t\}$, $\sum_{k=1}^{P_n} \psi_k^2 = O(P_n)$, $\max_{1 \leq k \leq P_n} |\psi_k| = O(1)$, $\min_{1 \leq k \leq P_n} |\psi_k \sigma^2| \geq 1$, $\nu > 0$, $\gamma_n$ and $P_n \to \infty$ with $\gamma_n / P_n^{1/2} = o(1/\ln n)$ and $P_n = O\left( (n/\gamma_n)^{1/14} \right) \leq \overline{p}_n/2$ and $\gamma_n \asymp (2 \ln \ln n)^{1/2}$ satisfies (3.1). We assume that $J_n = O\left(n^{1/2}\right)$.

**Proposition 1.** *Let $u_t$ be observed. Suppose that Assumptions K and P hold. For $\nu$ large enough, the alternative $G_1$ as above satisfies (3.4) and* *(i) the test (2.7) and its $\widehat{S}^{*}_{\widehat{p}^{*}}$ version consistently detect $G_1$. By contrast,*

*(ii) statistics $W_n$, $CvM_n$ and $EL_n$ have the same asymptotic distribution under $G_0$ and $G_1$ and the corresponding tests are therefore not consistent.*

Proposition 1-(ii) implies that tests based on $W_n$, $CvM_n$ or $EL_n$ are not adaptive rate-optimal. Let $\widehat{R}_{0,j}/\widehat{\tau}_{0,j}$ and $\widehat{R}_{1,j}/\widehat{\tau}_{1,j}$ be the standardized sample covariance computed under $G_0$ and $G_1$ respectively. It is established in the proof of Proposition 1 that

$$\max_{j\in[1,J_n]}\left|\frac{\widehat{R}_{0,j}}{\widehat{\tau}_{0,j}}-\frac{\widehat{R}_{1,j}}{\widehat{\tau}_{1,j}}\right|=o_{\mathbb{P}}\left(\frac{1}{(n\log n)^{1/2}}\right), \tag{4.5}$$

which implies that tests based on $W_n$ and $CvM_n$ are not consistent. The case of $EL_n$ test is a bit more involved but, due to its penalty scheme, this test statistic is asymptotically equal to $\widehat{BP}^{*}_{1}$ under the null and the alternative so that it cannot detect $G_1$ by (4.5).

## 5. Simulation experiments

Our simulation experiments aim to propose a valid penalty sequence $\gamma_n$ to be tested under various strong and weak white noise processes and under various alternatives. Since preliminary experiments have shown that the test statistic $\widehat{S}_{\widehat{p}}$ may yield an oversized test for some practically relevant white noise processes, we consider the test based on $\widehat{S}^{*}_{\widehat{p}^{*}}$ as in (2.8) and (2.9). To investigate the impact of choosing a large $\overline{p}_n$ latter on we allow for all possible orders, setting $\overline{p}_n=n-1$. We consider two kernels. The first is $K(t)=\mathbb{I}(t\in[0,1])$ which gives the Box-Pierce statistic so that the corresponding tests are labelled $BP$. The second uses the Parzen kernel

$$\mathsf{k}(t)=\begin{cases}1-6t^2+6|t|^3, & |t|\leq 1/2,\\ 2(1-|t|)^3, & 1/2<|t|\leq 1,\\ 0 & \text{otherwise.}\end{cases}$$

However since $\mathsf{k}(1)=0$ which would give a meaningless $\widehat{S}^{*}_{1}=0$, we change $\mathsf{k}(t)$ into $K(t)=\mathsf{k}(t/2)/\mathsf{k}(1/2)$ and label the corresponding tests as $Parz$. The critical values (2.10)

$\widehat{z}^*(\alpha)$, see also (2.5) and (2.6), use a power Parzen kernel $k(t) = \mathsf{k}^{32}(t)$, where the exponent 32 is has been proposed by Lee (2007) whose simulations show that such a choice ensures that the test with rejection region $n\widehat{R}_1^2 \geq \widehat{z}^*(\alpha)$ has good power properties. We consider 10%, 5% and 1% significance levels. A preliminary simulation experiment with $100,000$ replications gives that the corresponding quantiles $z_L(\alpha)$ of (2.4) used in $\widehat{z}^*(\alpha)$ are approximately 3.73, 5.58 and 10.97 respectively, which are in line with the critical values tabulated by Phillips et al. (2006, Table 6).

The first experiment analyzes the sensitivity of the test to the penalty term and aims to calibrate the proportionality constant for the penalty sequence. The experiment investigates the behavior of the test under the null for $\gamma_n = \gamma(2\ln\ln(n-2))^{1/2}$ where the proportionality coefficient $\gamma$ ranges from 2.8 to 3.8. The process $u_t$ is a white noise with the standard normal distribution. The next table reports the simulated levels for $50,000$ replications and the percentage $\%\{\widehat{p}^* \neq 1\}$ of simulation draws for which $\widehat{p}^* \neq 1$, an important indicator in deciding whether a difference between nominal and observed levels is due to a too small $\gamma_n$ or improper critical values. In Table 1, '*' indicates an oversized test, i.e. such that the null of a level smaller than the nominal size is rejected at 1% level by the one-sided test using the simulated level.

**[INSERT TABLE 1 HERE]**

A threshold value for the $BP$ test is $\gamma = 3.4$ which ensures that the observed sizes are close to the nominal sizes for $n = 1,000$. The $Parz$ test is slightly less oversized. Both tests have very similar value of $\%\{\widehat{p}^* \neq 1\}$, well below 1% for $\gamma = 3.4$. In the remaining simulation experiments $\gamma = 3.4$ is used.

We introduce some benchmark tests. We compare our $BP$ and $Parz$ tests with the data-driven test $EL$ based on the statistic $EL_n$ in (4.4) with $J_n = n - 1$ and the critical values of Lee (2007) in (2.10). We also consider the Newey-West data-driven order $\widehat{p}_{IMSE}$ used by

Hong and Lee (2005) and the test statistic

$$\widehat{p}_{IMSE} = \left(1 \vee \widehat{C}^{1/5}(f)\right) n^{1/5}, \quad \text{where} \quad \widehat{C}(f) = \frac{144 \sum_{j=-(n-1)}^{n-1} \mathsf{k}(j/\widetilde{p})\, j^4 \widehat{R}_j^2 / \widehat{\tau}_j^2}{0.539285 \sum_{j=-(n-1)}^{n-1} \mathsf{k}(j/\widetilde{p})\, \widehat{R}_j^2 / \widehat{\tau}_j^2},$$

$$IMSE = \frac{\sum_{j=1}^{\widehat{p}_{IMSE}} \mathsf{k}^2(j/\widehat{p}_{IMSE}) \left\{ \widehat{R}_j^2/\widehat{\tau}_j^2 - \left(1 - \frac{j}{n}\right) \right\}}{\left(2 \sum_{j=1}^{\widehat{p}_{IMSE}} \mathsf{k}^4(j/\widehat{p}_{IMSE}) \left(1 - \frac{j}{n}\right)^2\right)^{1/2}},$$

where $\mathsf{k}(\cdot)$ is the Parzen kernel and $\widehat{\tau}_j^2$ is defined as in (2.8). In the definition of $\widehat{p}_{IMSE}$, $\widetilde{p}$ is a pilot bandwidth that is set to $\widetilde{p} = 4(n/100)^{4/25}$. Note that $\widehat{C}(f)$ remains potentially stochastic under the null so that the null limit distribution of $IMSE$ may differ from the standard normal distribution valid for deterministic $p_n \to \infty$. We however follow common practice and use standard normal critical values for the $IMSE$ test. The last benchmark test, $CvM$, is based on Deo's (2000) Cramér-von Mises statistic $CvM_n$ in (4.3) and uses the critical values tabulated by Anderson and Darling (1952).

The first comparison under $\mathcal{H}_0$ is based on i.i.d. $\{u_t\}$ with the following distributions: standard normal ('Nor' in Table 2), Student with three degrees of freedom ('Stud'), and centered chi square with one degree of freedom ('Chi'). The Student distribution is used to test the sensitivity of our test to the lack of higher-order moments while the chi square distribution can reveal sensitivity to skewness.

**[INSERT TABLE 2 HERE]**

As in Table 1, the size of the $Parz$ test is slightly better than the size of the $BP$ test but both perform well here, although $BP$ is slightly oversized under the 'Chi' white noise. The $EL$ and $IMSE$ are generally oversized with strong size distortions for 'Chi'. The $CvM$ test performs well except for the 'Chi' experiment.

The next experiment considers observed weak white noise $u_t$ or residuals $\hat{u}_t$. Two conditional heteroskedastic martingale difference processes are examined. The first is a GARCH(1,1) process with $u_t = s_t\zeta_t$ and $s_t^2 = 0.001 + 0.90s_{t-1}^2 + 0.05u_{t-1}^2$ where $\zeta_t$ are i.i.d. standard normal innovations. The second process is an ARCH(1) process with $u_t = s_t\zeta_t$ and $s_t^2 = 0.001+0.9u_{t-1}^2$. Due to the ARCH coefficient larger than $\sqrt{1/3} = 0.577$, $\mathbb{E}\left[u_t^4\right] = \infty$ and the tests are, in principle, not expected to behave well in this experiment. The next three processes are uncorrelated but are not martingale differences, so that the $CvM$ test is not expected to have a correct size and is only reported here as a benchmark. The first, labelled 'Bilinear' in Table 3 below, is a bilinear model $u_t = \zeta_t + 0.9\zeta_{t-1}u_{t-2}$. The second, labelled 'No-MDS', is given by $u_t = \zeta_{t-1}\zeta_{t-2}\left(1 + \zeta_{t-2} + \zeta_t\right)$ and has been examined by Lobato (2001). The third, 'All-Pass', is an All-Pass ARMA(1,1) process examined by Lobato, Nankervis and Savin (2002), $u_t - 0.5u_{t-1} = \zeta_t - \zeta_{t-1}/0.5$, where $\zeta_t$ i.i.d. and have the Student distribution with 9 degrees of freedom. Since the root of the $MA$ part is the inverse of the $AR$ root, the resulting process is uncorrelated but the $u_t$ are dependent due to non-Gaussian $\zeta_t$. Finally, experiment 'ARRes' examines residuals from the $AR\left(1\right)$ $y_t = 0.8y_{t-1} + u_t$, $\widehat{u}_t = y_t - \widehat{\theta}y_{t-1}$, $\widehat{\theta} = \sum_{t=0}^{n-1} y_t y_{t+1} / \sum_{t=0}^{n-1} y_t^2$. The $BP$, $Parz$ and $EL$ tests are all adapted to the estimation effect thanks to the use of the critical values $\widehat{z}^*\left(\alpha\right)$ of (2.10). The critical values of the $IMSE$ and $CvM$ do not account for estimation of residuals and the corresponding tests should be not be expected to have a correct level under 'ARRes'.

**[INSERT TABLE 3 HERE]**

The performance of the $BP$ and $Parz$ tests is very good with levels that are not oversized in general. However the $BP$ and $Parz$ tests can be undersized, see the case of 'ARCH(1)'. But even in this case the value of $\%\left\{\widehat{p}^* \neq 1\right\}$ remains very small suggesting that the size distortion is due to the critical values of Lee (2007).[2] The behavior of the $EL$ test is more erratic, with levels that can be either oversized, as in the case of 'GARCH(1,1)', 'All Pass'

[2]This is confirmed by a not reported simulation experiment which shows that using standard chi-squared critical values give good results.

and 'ARRes', or undersized. The $IMSE$ test can also be severely oversized. The $CvM$ behaves well for 'GARCH(1,1)' and 'ARCH(1)' but, as expected, is severely size distorted in the other cases.

We now consider $\mathcal{H}_1$. In what follows, the critical values of the $EL$ and $IMSE$ tests are adjusted to achieve the desired level under normality. A first set of fixed alternatives is considered, $MA1$: $u_t = \varepsilon_t + 0.05\varepsilon_{t-1}$, $AR1$: $u_t = 0.05u_{t-1} + \varepsilon_t$, $MA4$: $\varepsilon_t + 0.2\varepsilon_{t-4}$ and $AR6$: $u_t = 0.3u_{t-6} + \varepsilon_t$ with i.i.d. standard normal innovations $\varepsilon_t$ and $n = 200,\ 1,000$ is considered. The $CvM$ test is expected to perform better for these alternatives, especially '$AR1$' and '$MA1$'. In Tables 4 and 5, $\overline{\widehat{p}^*}$ and $s_{\widehat{p}^*}$ are the simulation mean and standard deviation of $\widehat{p}^*$. These statistics are useful for assessing the impact of $\overline{p}_n$ on the power since large $\overline{\widehat{p}^*}$ or $s_{\widehat{p}^*}$ suggests that decreasing $\overline{p}_n$ can decrease the power.

**[INSERT TABLE 4 HERE]**

The low-lag '$AR1$' and '$MA1$' experiments have very similar characteristics with powers of the tests for $\alpha = 10\%$ increasing from $17\% - 18\%$ for $n = 200$ to $43\% - 47\%$ for $n = 1,000$. The data-driven tests all exhibit a surprisingly high $\overline{\widehat{p}^*}$ or $s_{\widehat{p}^*}$. The $BP$, $Parz$ and $EL$ seem to be outperformed by the $IMSE$ and $CvM$ tests. For the higher-order experiments '$MA4$' and '$AR6$' and $n = 1,000$, the $BP$, $Parz$ and $EL$ tests clearly outperform their competitors with power close or equal to 100%. For $n = 200$, the $EL$ test outperforms its competitors with $BP$ as a second-best. The high values of $\overline{\widehat{p}^*}$ and $s_{\widehat{p}^*}$ for the $BP$ and $Parz$ tests illustrate the fact that $\widehat{p}^*$ is suitable for testing but not as an estimator of the order of an $AR$ or $MA$ process.

The second experiment under $\mathcal{H}_1$ examines, for $n = 200$, the power of the 5% level $BP$ and $Parz$ tests against $H_\rho : u_t = v_t - \rho v_{t-1}$, $\rho \in [0, 1/2]$, under the nine scenarios of Tables 2 and 3. For example, under 'GARCH(1,1)' $v_t = s_t\zeta_t$ and $s_t^2 = 0.001 + 0.90s_{t-1}^2 + 0.05v_{t-1}^2$ where $\zeta_t$ are i.i.d. standard normal innovations while, under 'ARRes', the $v_t$ are i.i.d. $N(0, 1)$

and $u_t = v_t - \rho v_{t-1}$ is estimated from the $AR(1)$ model $X_t = 0.8X_{t-1} + u_t$. We do not consider the other tests to avoid undesirable size correction effects, but we compare $BP$ and $Parz$ with $\widetilde{\mathcal{M}}_n^{EP32}$ test of Lee (2007) which rejects the null when $n\widehat{R}_1^2 \geq \widehat{z}(\alpha)$ where $\widehat{z}(\alpha)$ is defined in (2.7), and an $\alpha$ level test which rejects the null when $n\widehat{R}_1^2 \geq c(\alpha)$, where the infeasible $c(\alpha)$, dependent of the white noise process under consideration, is computed from $10,000$ preliminary replications. Since the latter is locally optimal under Gaussianity, it is labelled $LOT$. Figure 1 reports the nine power graphs corresponding to each white noise experiments.

**[INSERT FIGURE 1 HERE]**

Except for white noise processes such as 'NoMDS' for which the new tests are undersized, the power of the four tests are quite similar in the vicinity of $\rho = 0$, suggesting that our data-driven tests are, for processes close to Gaussianity, not far from being locally optimal as $LOT$. The global performance of all tests deteriorate for nonlinear white noise processes as 'ARCH(1)', for which $LOT$ has a very low power compared to its competitors $BP$, $Parz$ and $\widetilde{\mathcal{M}}_n^{EP32}$. $Parz$ dominates its competitors for such white noise processes. As expected from (3.8), $Parz$ and $BP$ perform as well as or better than $\widetilde{\mathcal{M}}_n^{EP32}$ which is less powerful than $Parz$ for heteroskedastic noises the 'Bilinear', 'ARCH(1)', 'GARCH(1,1)' or 'NoMDS'.

The third experiment under $\mathcal{H}_1$ considers a second set of alternatives given by randomized "small correlation" processes defined in (4.1),

$$u_t = \varepsilon_t + \frac{(2.5 \times \gamma_n)^{1/2}}{n^{1/2} P^{1/4}} \sum_{k=1}^{P} \psi_{k,b} \varepsilon_{t-k}, \qquad \psi_{k,b} \overset{\text{i.i.d.}}{\sim} N(0,1). \tag{5.1}$$

In this setting $b$ is the simulation index, $b = 1, ..., 10,000$. New moving average coefficients $\{\psi_{k,b}\}$ are drawn for each simulation. Randomizing the moving average coefficients allows us to explore various shapes of the correlation function. The noise $\{\varepsilon_t\}$ is independent of the moving average coefficients $\{\psi_{k,b}\}$ and is drawn randomly from the standard normal distribution. Since $\sum_{k=1}^{P} \psi_{k,b}^2 = P(1 + o_{\mathbb{P}}(1))$ when $P$ tends to infinity, the covariances of

(5.1) can be $o\left(n^{-1/2}\right)$ as shown in Lemma 1. We consider two scenarios. In the experiment 'LOW', $P$ is set to 15 for $n = 200$ and to 75 when $n = 1,000$. The experiment 'HIGH' doubles the order $P$, so $P = 30$ for $n = 200$ and $P = 150$ for $n = 1,000$. The next table reports simulation results.

**[INSERT TABLE 5 HERE]**

The $BP$ test outperforms its competitors and $Parz$ comes as a second-best. The $EL$ test achieves power similar to that of the $BP$ test only in the LOW experiment when $P = 15$ and $n = 200$. The power of the $IMSE$ and $CvM$ tests decreases with the sample size while the power of the other tests increases, showing the importance of a proper data-driven choice of the order. The high values of $\widehat{\overline{p}^{*}}_{Parz}$ may suggest that the $Parz$ test would be negatively affected by choosing a lower value of $\overline{p}_n$. However setting $\overline{p}_n = 3\left[(n/2)^{1/2}\right]$ instead of $\overline{p}_n = n-1$ as done in an experiment not reported does not really affect the power of the $BP$ test.

## 6. Concluding remarks

The paper proposes an automatic test for the weak white noise null hypothesis for observed variables or residuals from a parametric model. The test is based on a new data-driven order selection procedure applied to the Box-Pierce (1970) test statistic. The critical region uses robust critical values of Lee (2007) which can account for estimation of residuals. An important theoretical finding is that the new test can detect alternatives with small autocorrelation coefficients of order $\rho_n = o\left(n^{-1/2}\right)$ where $n$ is the sample size, provided that the number of autocorrelation coefficients at moderate lags is large enough. The proposed test is shown to be adaptive rate-optimal against this class of alternatives. The paper gives examples of moving average alternatives with small autocorrelation coefficients of order $o\left(n^{-1/2}\right)$ which are detected by the new test but not by tests previously proposed by Deo (2000), Escanciano and Lobato (2009) or Xiao and Wu (2011). These alternatives correspond to a plausible macroeconomic scenario where a temporary shock has no significant impact whereas permanent shocks may cause significant changes. They can also be of interest in finance where arbitrage should rule out strong deviations from the difference of martingale hypothesis, since these alternatives generate conditional expectation given the past of these alternatives with order $o_{\mathbb{P}}\left(n^{-1/2}\right)$. A simulation experiment has shown that the new test can cope with various weak types of white noise processes including the ARCH or GARCH processes popular in empirical finance. The simulation experiment has also confirmed good power properties of the test regarding detection of standard $AR(1)$ and $MA(1)$ alternatives when the noise is highly nonlinear, for instance in the case of the $ARCH(1)$ process considered in the experiment.

## 7. References


Anderson, T.W. (1993). Goodness of Fit Tests for Spectral Distributions. *The Annals of Statistics* **21**, 830–847.

Anderson, T.W. and D.A. Darling (1952). Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *Annals of Mathematical Statistics* **23**, 193–212.

Box, G. and D. Pierce (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of American Statistical Association* **65**, 1509–1526.

Campbell, J.Y., A.W. Lo and A.C. Craig MacKinlay (1997). *The Econometrics of Financial Markets*. Second Edition, Princeton University Press.

Chen, S.X. and J. Gao (2007). An adaptive Empirical Likelihood Test for Parametric Time Series Regression Models. *Journal of Econometrics* **141**, 950–972.

Delgado, M.A. and C. Velasco (2012). An Asymptotically Pivotal Transform of the Residuals Sample Autocorrelations with Application to Model Checking. *Journal of American Statistical Association* **106**, 646—958.

Delgado, M.A., J. Hidalgo and C. Velasco (2005). Distribution Free Goodness-of-Fit Tests for Linear Processes. *The Annals of Statistics* **33**, 2568-2609.

Deo, R.S. (2000). Spectral Tests of the Martingale Hypothesis under Conditional Heteroscedasticity. *Journal of Econometrics* **99**, 291-315.

Donoho, D. and J. Jin (2004). Higher Criticism for Detecting Sparse Heterogeneous Mixtures. *The Annals of Statistics* **32**, 962–994.

Durlauf, S.N. (1991). Spectral Based Testing of the Martingale Hypothesis. *Journal of Econometrics* **50**, 355-376.

Escanciano, J.C. and I.N. Lobato (2009). An Automatic Portmanteau Test for Serial Correlation. *Journal of Econometrics* **151**, 140–149.

Fan, J. (1996). Test of Significance Based on Wavelet Thresholding and Neyman's Truncation. *Journal of the American Statistical Association* **91**, 674–688.

Fan, J. and Q. Yao (2005). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer.

Francq, C. , R. Roy and J.M. Zakoian (2005). Diagnostic Checking in ARMA Models With Uncorrelated Errors. *Journal of the American Statistical Association* **100**, 532–544.

Golubev, G.K., M. Nussbaum and H.H. Zhou (2010). Asymptotic Equivalence of Spectrum Density Estimation and Gaussian White Noise. *The Annals of Statistics* **38**, 181–214.

Grenander, U. and M. Rosenblatt (1952). On Spectral Analysis of Stationary Time-series. *Proceedings of the National Academy of Sciences U.S.A.* **38,** 519-521.

Guay, A. and E. Guerre (2006). A Data-Driven Nonparametric Specification Test for Dynamic Regression Models. *Econometric Theory* **22**, 543–586.

Guerre, E. and P. Lavergne (2002). Optimal Minimax Rates for Nonparametric Specification Testing in Regression Models. *Econometric Theory* **18**, 1139–1171.

Guerre, E. and P. Lavergne (2005). Rate-Optimal Data-Driven Specification Testing for Regression Models. *The Annals of Statistics* **33**, 840–870.

Hong, Y. (1996). Consistent Testing for Serial Correlation of Unknown Form. *Econometrica* **64**, 837–864.

Hong, Y. and Y.J. Lee. (2005). Generalized Spectral Tests for Conditional Mean Models in Time Series with Conditional Heteroscedasticity of Unknown Form. *Review of Economic Studies* **72**, 499–541.

Horowitz, J.L. and V.G. Spokoiny (2001). An Adaptive, Rate-Optimal Test of a Parametric Mean-Regression Model Against a Nonparametric Alternative. *Econometrica* **69**, 599–631.

Kuan, C.M. and W.M. Lee (2006). Robust $M$ Tests without Consistent Estimation of the Asymptotic Covariance Matrix. *Journal of the American Statistical Association* **101**, 1264–1275.

Lee, W.M. (2007). Robust M Tests Using Kernel-based Estimators with Bandwidth Equal to Sample Size. *Economics Letters* **96**, 295–300.

Lobato, I.N. (2001). Testing That a Dependent Process Is Uncorrelated. *Journal of the American Statistical Association* **96**, 1066–1076.

Lobato, I.N., J.C. Nankervis and N.E. Savin (2002). Testing for Zero Autocorrelation in the Presence of Statistical Dependence. *Econometric Theory* **18**, 730–743.

Newey, W.K. and K. West (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies* **61**, 631–653.

Phillips, P.C.B, Y. Sun & S. Jin (2006). Spectral Density Estimation and Robust Hypothesis Testing Using Steep Origin Kernels Without Truncation. *International Economic Review* **47**, 837–894.

Pollard, D. (2002). *A User's Guide to Measure Theoretic Probability*. Cambridge University Press.

Shao, X. (2011a). A Bootstrap-assisted Spectral Test of White Noise under Unknown Dependence. *Journal of Econometrics* **162**, 213–224.

Shao, X. (2011b). Testing for White Noise under Unknown Dependence and its Applications to Goodness-of-Fit for Time Series Models. *Econometric Theory* **27**, 312–343.

Spokoiny, V.G. (1996). Adaptive Hypothesis Testing Using Wavelets. *The Annals of Statistics* **24**, 2477–2498.

Xiao, H. and W.B. Wu (2011). Asymptotic Inference of Autocovariances of Stationary Processes. University of Chicago, arXiv:11053423v1.

Wu, W.B. (2005). Nonlinear System Theory: Another Look at Dependence. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 14150–14154.

Wu, W.B. (2007). Strong Invariance Principles for Dependent Random Variables. *The Annals of Probability* **35**, 2294–2320.

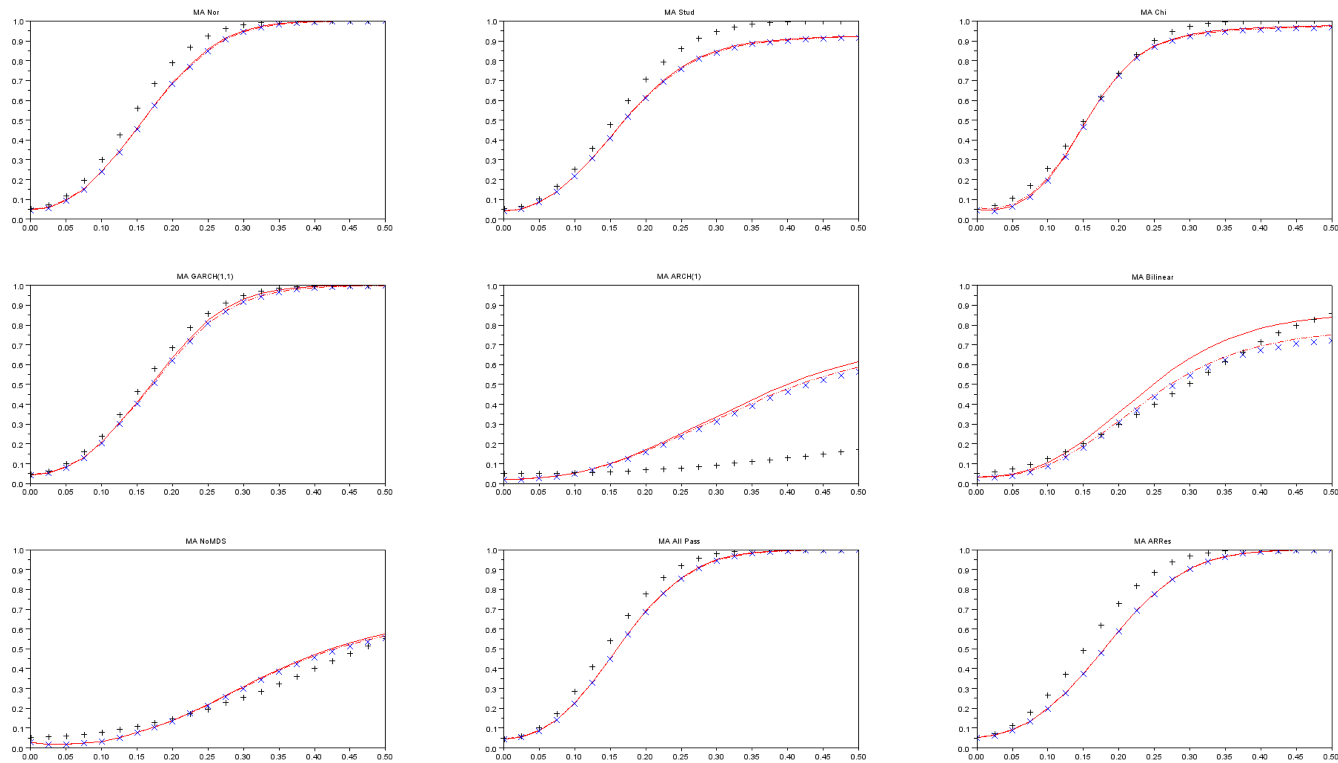


FIGURE 1. **Empirical rejection probabilities** of $LOT$ (black '+' line), $Parz$ (red solid line), $BP$ (red dotted line) and Lee (2007) $\widetilde{\mathcal{M}}^{EP32}$ test (blue 'x' line). The level of these tests is 5% . The alternative is an $MA(1)$ with a moving average coefficient ranging from 0 to 1/2 and disturbances as in Tables 2-3. The sample size is $n = 200$ and the number of replications is $10,000$.

| $\gamma$ | 2.8 | | 3.0 | | 3.2 | | 3.4 | | 3.6 | | 3.8 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 | 200 | 1.000 |
| $\alpha_{BP} = 10\%$ | 10.91* | 10.47* | 10.52* | 10.30* | 10.33* | 10.15 | 10.18 | 10.06 | 10.08 | 10.06 | 9.96 | 9.98 |
| $\alpha_{BP} = 5\%$ | 6.06* | 5.50* | 5.65* | 5.31* | 5.45* | 5.15* | 5.30* | 5.07 | 5.19 | 5.02 | 5.07 | 4.98 |
| $\alpha_{BP} = 1\%$ | 1.95* | 1.46* | 1.57* | 1.29* | 1.39* | 1.15 | 1.25* | 1.07 | 1.16* | 1.04 | 1.04 | 1.00 |
| $\%\{\widehat{p}^*_{BP} \neq 1\}$ | 1.37 | 0.65 | 0.91 | 0.44 | 0.69 | 0.28 | 0.53 | 0.19 | 0.41 | 0.14 | 0.27 | 0.10 |
| $\alpha_{Parz} = 10\%$ | 10.02 | 9.97 | 9.93 | 9.94 | 9.88 | 9.91 | 9.83 | 9.90 | 9.81 | 9.90 | 9.79 | 9.89 |
| $\alpha_{Parz} = 5\%$ | 5.18* | 4.99 | 5.06 | 4.94 | 5.00 | 4.92 | 4.95 | 4.90 | 4.92 | 4.90 | 4.90 | 4.90 |
| $\alpha_{Parz} = 1\%$ | 1.43* | 1.18* | 1.23* | 1.09* | 1.10* | 1.02 | 1.01* | 0.99 | 0.95 | 0.97 | 0.91 | 0.95 |
| $\%\{\widehat{p}^*_{Parz} \neq 1\}$ | 0.86 | 0.43 | 0.65 | 0.31 | 0.48 | 0.22 | 0.36 | 0.11 | 0.28 | 0.07 | 0.20 | 0.10 |

Table 1. **Penalty sequence impact on levels ($\gamma_n = \gamma\,(2\ln\ln(n-2))^{1/2}$, 50,000 replications).** A "*" indicates an oversized test at the 1% level.

| | Tests | BP | | Parz | | EL | | IMSE | | CvM | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\{u_t\}$ | $n$ | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 |
| Nor | $\alpha = 10\%$ | 10.37 | 10.21 | 10.06 | 10.10 | 12.30* | 10.92 | 11.46* | 10.70* | 9.79 | 9.54 |
| | $\alpha = 5\%$ | 5.28 | 4.85 | 4.98 | 4.73 | 7.06* | 5.56* | 7.80* | 7.54* | 4.60 | 4.88 |
| | $\alpha = 1\%$ | 1.12 | 0.96 | 0.94 | 0.91 | 2.06* | 1.43* | 4.19* | 3.80* | 1.01 | 0.88 |
| | $\%\{\widehat{p}^* \neq 1\}$ | 0.48 | 0.14 | 0.28 | 0.11 | 2.61 | 0.94 | | | | |
| Stud | $\alpha = 10\%$ | 9.33 | 10.38 | 9.16 | 10.25 | 10.88* | 10.76 | 11.07* | 11.20* | 9.76 | 10.02 |
| | $\alpha = 5\%$ | 4.32 | 4.93 | 4.17 | 4.80 | 5.74* | 5.30 | 7.89* | 7.57* | 4.63 | 4.88 |
| | $\alpha = 1\%$ | 0.80 | 0.87 | 0.72 | 0.79 | 1.44* | 1.13 | 3.96* | 3.71* | 0.81 | 0.95 |
| | $\%\{\widehat{p}^* \neq 1\}$ | 0.27 | 0.15 | 0.21 | 0.10 | 2.12 | 0.58 | | | | |
| Chi | $\alpha = 10\%$ | 11.00* | 10.92* | 10.10 | 10.28 | 14.15* | 12.30* | 14.54* | 13.11* | 12.02* | 11.27* |
| | $\alpha = 5\%$ | 5.61* | 5.83* | 4.70 | 5.17 | 8.70* | 7.18* | 10.37* | 9.53* | 6.32* | 5.76* |
| | $\alpha = 1\%$ | 1.83* | 1.72* | 1.08* | 1.12* | 3.91* | 2.64* | 5.97* | 5.37* | 1.46* | 1.48* |
| | $\%\{\widehat{p}^* \neq 1\}$ | 1.19 | 0.84 | 0.60 | 0.39 | 4.71 | 2.49 | | | | |

Table 2. **i.i.d distributions ($\gamma_n = 3.4\,(2\ln\ln(n-2))^{1/2}$, 10,000 replications).** A "*" indicates an oversized test at the 1% level.

| $\{u_t\}$ | Tests / $n$ | BP 200 | BP 1,000 | Parz 200 | Parz 1,000 | EL 200 | EL 1,000 | IMSE 200 | IMSE 1,000 | CvM 200 | CvM 1,000 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GARCH(1,1) | $\alpha = 10\%$ | 9.93 | 10.00 | 9.56 | 9.88 | 11.78* | 10.75 | 11.08* | 10.88* | 9.38 | 9.56 |
| | $\alpha = 5\%$ | 5.02 | 4.75 | 4.66 | 4.63 | 6.69* | 5.49* | 7.70* | 7.72* | 4.64 | 4.82 |
| | $\alpha = 1\%$ | 1.03 | 0.77 | 0.78 | 0.73 | 1.99* | 1.25* | 4.20* | 3.85* | 0.94 | 0.93 |
| | $\%\{\widehat{p}^* \neq 1\}$ | 0.54 | 0.13 | 0.24 | 0.11 | 2.64 | 0.99 | | | | |
| ARCH(1) | $\alpha = 10\%$ | 6.01 | 6.46 | 5.96 | 6.42 | 7.04 | 6.78 | 10.39 | 10.03 | 9.51 | 9.56 |
| | $\alpha = 5\%$ | 2.33 | 2.44 | 2.30 | 2.42 | 3.31 | 2.68 | 7.06* | 6.94* | 4.26 | 4.23 |
| | $\alpha = 1\%$ | 0.35 | 0.22 | 0.34 | 0.24 | 1.09 | 0.37 | 3.61* | 3.42* | 0.65 | 0.72 |
| | $\%\{\widehat{p}^* \neq 1\}$ | 0.22 | 0.09 | 0.20 | 0.11 | 1.61 | 0.47 | | | | |
| Bilinear | $\alpha = 10\%$ | 8.03 | 8.42 | 7.88 | 8.35 | 9.36 | 8.87 | 14.52* | 14.22* | 13.02* | 12.99* |
| | $\alpha = 5\%$ | 3.34 | 3.46 | 3.30 | 3.41 | 4.51 | 3.91 | 10.60* | 10.50* | 7.08* | 6.93* |
| | $\alpha = 1\%$ | 0.67 | 0.60 | 0.81 | 0.64 | 1.46 | 0.83 | 6.02* | 5.61* | 1.49* | 1.69* |
| | $\%\{\widehat{p}^* \neq 1\}$ | 0.47 | 0.22 | 0.58 | 0.33 | 2.23 | 0.75 | | | | |
| NoMDS | $\alpha = 10\%$ | 8.13 | 9.95 | 8.13 | 9.89 | 9.51 | 10.53 | 14.38* | 16.08* | 13.55* | 12.99* |
| | $\alpha = 5\%$ | 2.92 | 4.19 | 2.96 | 4.15 | 4.08 | 4.71 | 9.69* | 11.23* | 6.56* | 8.02* |
| | $\alpha = 1\%$ | 0.31 | 0.54 | 0.44 | 0.62 | 0.93 | 0.84 | 5.10* | 5.92* | 1.10 | 1.67 |
| | $\%\{\widehat{p}^* \neq 1\}$ | 0.09 | 0.09 | 0.21 | 0.23 | 1.84 | 0.78 | | | | |
| All Pass | $\alpha = 10\%$ | 10.07 | 10.06 | 9.86 | 10.00 | 11.71* | 10.70 | 8.35 | 8.44 | 6.93 | 6.73 |
| | $\alpha = 5\%$ | 4.76 | 5.20 | 4.53 | 5.14 | 6.42* | 5.84 | 5.64* | 5.54 | 2.99 | 3.02 |
| | $\alpha = 1\%$ | 1.04* | 1.08 | 0.84 | 1.04 | 2.06* | 1.52* | 2.57* | 2.19* | 0.43 | 0.41 |
| | $\%\{\widehat{p}^* \neq 1\}$ | 0.35 | 0.10 | 0.16 | 0.09 | 2.23 | 0.80 | | | | |
| ARRes | $\alpha = 10\%$ | 10.07 | 10.36 | 9.92 | 10.19 | 10.90* | 11.17 | 4.07 | 4.10 | 3.98 | 4.05 |
| | $\alpha = 5\%$ | 5.16 | 5.16 | 5.00 | 4.97 | 5.98* | 6.08 | 2.84 | 2.76 | 1.80 | 1.73 |
| | $\alpha = 1\%$ | 1.26 | 1.30 | 1.09 | 1.06 | 2.00* | 2.06 | 1.26 | 1.32 | 0.19 | 0.23 |
| | $\%\{\widehat{p}^* \neq 1\}$ | 0.27 | 0.30 | 0.13 | 0.08 | 1.30 | 1.36 | | | | |

TABLE 3. **Weak white noise and estimated residuals** ($\gamma_n = 3.4\,(2\ln\ln(n-2))^{1/2}$, 10,000 replications)**.** A "*" indicates an oversized test at the 1% level.

| | Tests | BP | | Parz | | EL$^{esc}$ | | IMSE$^{esc}$ | | CvM | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\{u_t\}$ | $n$ | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 |
| MA1 | $\alpha = 10\%$ | 17.6 | 43.6 | 17.3 | 43.5 | 16.25 | 42.9 | 17.3 | 42.7 | 17.83 | 46.6 |
| | $\alpha = 5\%$ | 9.7 | 29.9 | 9.4 | 29.3 | 8.15 | 27.89 | 10.58 | 31.03 | 10.76 | 34.4 |
| | $\alpha = 1\%$ | 2.5 | 10.0 | 2.5 | 11.0 | 1.5 | 8.0 | 2.7 | 13.4 | 2.8 | 15.1 |
| | $\overline{\widehat{p}^*}$ $(s_{\widehat{p}^*})$ | 1.3 (6) | 1.2 (12) | 1.3 (8) | 1.2 (1) | 1.1 (1) | 1.0 (1) | 5.7 (1) | 7.0 (1) | | |
| AR1 | $\alpha = 10\%$ | 17.4 | 43.6 | 17.0 | 43.5 | 15.9 | 42.1 | 17.1 | 43.1 | 17.9 | 46.5 |
| | $\alpha = 5\%$ | 10.0 | 30.4 | 9.7 | 30.4 | 8.5 | 28.8 | 10.4 | 31.6 | 10.5 | 35.1 |
| | $\alpha = 1\%$ | 2.6 | 11.2 | 2.7 | 12.4 | 1.6 | 8.6 | 2.5 | 14.2 | 2.7 | 8.6 |
| | $\overline{\widehat{p}^*}$ $(s_{\widehat{p}^*})$ | 1.3 (6) | 1.0 (4) | 1.3 (7) | 1.2 (1) | 1.0 (1) | 1.0 (1) | 5.7 (1) | 7.0 (1) | | |
| MA4 | $\alpha = 10\%$ | 17.1 | 97.6 | 11.9 | 81.5 | 23.4 | 98.3 | 14.1 | 99.4 | 13.5 | 50.8 |
| | $\alpha = 5\%$ | 12.6 | 97.5 | 7.3 | 80.9 | 19.4 | 98.2 | 8.0 | 98.1 | 7.0 | 23.4 |
| | $\alpha = 1\%$ | 8.9 | 97.1 | 3.7 | 80.4 | 12.9 | 97.7 | 1.6 | 89.4 | 1.6 | 4.5 |
| | $\overline{\widehat{p}^*}$ $(s_{\widehat{p}^*})$ | 3.3 (16) | 13.1 (80) | 3.5 (19) | 35.2 (91) | 1.9 (3) | 5.4 (4) | 7.0 (1) | 12.5 (1) | | |
| AR6 | $\alpha = 10\%$ | 46.3 | 100 | 28.5 | 100 | 71.1 | 100 | 17.8 | 26.3 | 19.1 | 64.0 |
| | $\alpha = 5\%$ | 43.1 | 100 | 24.8 | 100 | 70.0 | 100 | 10.7 | 17.3 | 10.7 | 32.2 |
| | $\alpha = 1\%$ | 40.8 | 100 | 22.2 | 100 | 61.6 | 100 | 3.0 | 7.1 | 2.7 | 6.5 |
| | $\overline{\widehat{p}^*}$ $(s_{\widehat{p}^*})$ | 14.5 (37) | 158 (314) | 23.5 (53) | 204 (307) | 6.8 (7) | 12.9 (10) | 6.1 (1) | 7.4 (1) | | |

TABLE 4. **AR-MA alternatives** ($\gamma_n = 3.4\,(2\ln\ln(n-2))^{1/2}$, 10,000 replications). "*esc*" indicates empirical size-corrected power.

| | Tests | BP | | Parz | | EL$^{esc}$ | | IMSE$^{esc}$ | | CvM | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\{u_t\}$ | $n$ | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 | 200 | 1,000 |
| LOW | $\alpha = 10\%$ | 75.5 | 94.8 | 66.9 | 90.7 | 74.2 | 71.3 | 66.8 | 57.7 | 61.1 | 49.0 |
| | $\alpha = 5\%$ | 71.9 | 94.1 | 66.9 | 89.8 | 69.3 | 67.6 | 57.6 | 47.4 | 48.1 | 35.7 |
| | $\alpha = 1\%$ | 66.2 | 93.1 | 58.4 | 88.9 | 59.5 | 61.4 | 37.5 | 28.2 | 27.8 | 17.7 |
| | $\overline{\widehat{p}^*}\ (s_{\widehat{p}^*})$ | 40.2 (61) | 363 (369) | 67.6 (78) | 579 (367) | 9.9 (12) | 36.2 (34) | 7.5 (1) | 8.7 (1) | | |
| HIGH | $\alpha = 10\%$ | 73.8 | 95.6 | 67.0 | 93.1 | 67.8 | 61.7 | 57.5 | 48.1 | 51.2 | 39.9 |
| | $\alpha = 5\%$ | 70.3 | 95.6 | 63.3 | 93.1 | 62.2 | 57.1 | 47.2 | 36.8 | 39.8 | 27.9 |
| | $\alpha = 1\%$ | 65.5 | 94.7 | 59.1 | 92.0 | 52.1 | 49.7 | 27.8 | 19.5 | 20.9 | 12.3 |
| | $\overline{\widehat{p}^*}\ (s_{\widehat{p}^*})$ | 58.0 (68) | 557 (340) | 95.1 (91) | 851 (297) | 13.3 (16) | 42.4 (54) | 7.2 (1) | 8.4 (1) | | |

TABLE 5. **Small correlations alternatives ($\gamma_n = 3.4\,(2\ln\ln(n-2))^{1/2}$, 10,000 replications). "$esc$" indicates empirical size-corrected power.**

Robust Adaptive Rate-Optimal Testing for the White Noise Hypothesis: Supplementary Material

Alain Guay[3]

Emmanuel Guerre[4]

Štěpána Lazarová[5]

This version: 2nd November 2012

[3]CIRPÉE and CIREQ, Université du Québec à Montréal, e-mail: guay.alain@uqam.ca

[4]School of Economics and Finance, Queen Mary, University of London, e-mail: e.guerre@qmul.ac.uk

[5]School of Economics and Finance, Queen Mary, University of London, e-mail: s.lazarova@qmul.ac.uk

# Supplementary Material A: proofs of main results

This section contains the proofs of the results of Section 3. In what follows, a tilde superscript, as in

$$\widetilde{S}_p = n \sum_{j=1}^{p} K^2 \left(\frac{j}{p}\right) \widetilde{R}_j^2 \quad \text{where} \quad \widetilde{R}_j = \frac{1}{n} \sum_{t=1}^{n-|j|} u_t u_{t+|j|}. \tag{A.1}$$

indicates that the variables $u_t$ are observed. This also leads to define

$$\widetilde{\tau}_j = \frac{1}{n} \sum_{t=1}^{n-|j|} u_t^2 u_{t+|j|}^2, \quad \widetilde{z}_L(\alpha) = \widehat{z}_L(\alpha), \quad \widetilde{z}_L^*(\alpha) = \widehat{z}_L^*(\alpha),$$

but we keep the notation $\widehat{p}$. $C$ and $C'$ are constants that may vary from line to line but only depend on the constants of the assumptions. Notation $[\cdot]$ is used for the integer part of a real number and $a \vee b = \max(a,b)$, $a \wedge b = \min(a,b)$. Let $\overline{u}_t^{t-j} = \overline{u}_{t,n}^{t-j}$ be a copy of $u_t = F_n(\ldots, e_{t-1}, e_t)$ obtained by changing $e_{t-j}$, $e_{t-j-1}$, $\ldots$ into $e'_{t-j}$, $e'_{t-j-1}$, $\ldots$. Then the condition $\left\| u_t - u_t^{t-j} \right\|_a \leq \delta_a(j)$ ensures that

$$\left\| u_t - \overline{u}_t^{t-j} \right\|_a \leq \Theta_a(j) \text{ where } \Theta_a(j) = \sum_{i=j}^{\infty} \delta_a(j). \tag{A.2}$$

We first state some intermediary results that are used in the proofs of our main results. These intermediary results are proven in a section called "Supplementary Material B". Lemma A.2 gives the order of standardization terms $E(p)$, $E_\Delta(p)$ and $V_\Delta(p)$. Propositions A.1 and A.2 deal with the impact of the estimation of $\theta$. Proposition A.3 is used to study the asymptotic null behavior of the test and to show that $\mathbb{P}(\widehat{p} = 1) \to 1$ in Theorem 1. Proposition A.3 deals with observed variables or residuals thanks to Propositions A.1 and A.2. Propositions A.4 and A.5 are the key tools for our consistency result, Theorem 2. They dealt with observed variables but are combined with Propositions A.1 and A.2 to deal with estimation errors in the proof of Theorem 2.

**Lemma A.2.** *Suppose Assumption K holds and that $\overline{p}_n/n \leq 1/2$. (i) There exists a constant $C > 1$ such that, for $q = 1, 2$ and for any $1 \leq p \leq \overline{p}_n$, $\frac{p}{C} \leq \sum_{j=1}^{n-1} \left(1 - \frac{j}{n}\right)^q K^{2q}\left(\frac{j}{p}\right) \leq$*

$Cp$, $\frac{p}{C} \leq \sum_{j=1}^{n-1} K^{2q}\left(\frac{j}{p}\right) \leq Cp$, $V_\Delta^2(p) \leq Cp$, *and* $E_\Delta(p) \leq \sum_{j=1}^{n-1}\left(K^2\left(\frac{j}{p}\right) - K^2(j)\right) \leq Cp^{1/2}V_\Delta(p)$; *(ii) Under Assumption P, for all* $n$ *and all* $p \in [1, \overline{p}_n]$, $V_\Delta(p) \geq C(p-1)^{1/2}$ *and* $E_\Delta(p) \geq 0$.

**Lemma A.3.** *Suppose Assumptions K, M and R hold. Then the rejection regions* $\widetilde{S}_1 \geq \widetilde{z}_L(\alpha)$, $\widetilde{S}_1^* \geq \widetilde{z}_L^*(\alpha)$, $\widehat{S}_1 \geq \widehat{z}_{KL}(\alpha)$ *and* $\widehat{S}_1^* \geq \widehat{z}_{KL}^*(\alpha)$ *are asymptotically of level* $\alpha$*. Moreover, under* $\mathcal{H}_1$, $\widehat{z}_L(\alpha)$, $\widetilde{z}_L^*(\alpha)$, $\widehat{z}_{KL}(\alpha)$ *and* $\widehat{z}_{KL}^*(\alpha)$ *are all* $O_{\mathbb{P}}(1)$.

**Lemma A.4.** *Under Assumption R,* $\sup_{0\leq j\leq n-1} \operatorname{Var}\left(\widetilde{R}_j\right) \leq \frac{C}{n}$.

**Proposition A.1.** *Suppose Assumptions M, P and R hold. Then* $\max_{j\in[0,\overline{p}_n]}\left|\widehat{R}_j - \widetilde{R}_j\right| = O_{\mathbb{P}}\left(n^{-1/2}\right)$, $\max_{p\in[0,n-1]} n\sum_{j=1}^{p}\left(\widehat{R}_j - \widetilde{R}_j\right)^2 = O_{\mathbb{P}}(1)$, *and*

$$\max_{j\in[0,n-1]}\left|\widetilde{R}_j - \left(1-\frac{j}{n}\right)R_{j,n}\right| = O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right),$$

$$\max_{j\in[0,\overline{p}_n]}\left|\widehat{R}_j - R_{j,n}\right| = O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right),$$

$$\max_{j\in[0,n-1]}\left(1-\frac{j}{n}\right)\left|\widetilde{\tau}_j^2 - \tau_{j,n}^2\right| = O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right),$$

$$\max_{j\in[0,\overline{p}_n]}\left|\widehat{\tau}_j^2 - \tau_{j,n}^2\right| = O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right).$$

**Proposition A.2.** *Let Assumptions K, M, P and R hold. Let* $\widetilde{S}_p$ *be as in (A.1). Then*

$$\max_{p\in[2,\overline{p}_n]}\frac{\left|\left(\widehat{S}_p - \widehat{S}_1\right) - \left(\widetilde{S}_p - \widetilde{S}_1\right)\right|}{1+\left(n\sum_{j=1}^{p}R_{j,n}^2\right)^{1/2}} = O_{\mathbb{P}}(1)$$

*and for any* $p_n = O(n^{1/2})$, $\widehat{S}_{p_n} - \widetilde{S}_{p_n} = O_{\mathbb{P}}\left(1+\left(n\sum_{j=1}^{p_n}R_{j,n}^2\right)^{1/2}\right)$.

**Proposition A.3.** *Suppose Assumptions K, M, P and R hold and that* $\mathcal{H}_0$ *is true. Then (3.1) ensures that*

$$\lim_{n\to\infty}\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\frac{(\widehat{S}_p - \widehat{S}_1)/\widehat{R}_0^2 - E_\Delta(p)}{V_\Delta(p)} \geq \gamma_n\right) = 0.$$

**Proposition A.4.** *Under Assumptions K, P and R, there are some $C, C' > 0$ such that for $n$ large enough and uniformly in $p \in [1, \overline{p}_n]$,*

$$\mathbb{E}\left[\widetilde{S}_p\right] - R_{0,n}^2 E(p) \geq Cn \sum_{j=1}^{p/2} R_{j,n}^2 - C' R_{0,n}^2,$$

$$\mathbb{E}\left[\sum_{j=1}^{n-1} K\left(\frac{j}{p}\right) \frac{\widetilde{R}_j^2}{\tau_{j,n}^2}\right] - E(p) \geq Cn \sum_{j=1}^{p/2} \left(\frac{R_{j,n}}{R_{0,n}}\right)^2 - C'.$$

**Proposition A.5.** *Under Assumptions K, P and R, there is a constant $C > 0$ such that for $n$ large enough and uniformly in $p \in [1, \overline{p}_n]$,*

$$\operatorname{Var}\left(\widetilde{S}_p\right) \leq C\left(n \sum_{j=1}^{p} R_{j,n}^2 + p\right),$$

$$\operatorname{Var}\left(\sum_{j=1}^{n-1} K\left(\frac{j}{p}\right) \frac{\widetilde{R}_j^2}{\tau_{j,n}^2}\right) \leq C\left(n \sum_{j=1}^{p} \frac{R_{j,n}^2}{R_{0,n}^2} + p\right).$$

A.1. **Proof of Theorem 1.** (3.2), (3.1) and Proposition A.3 give that $\lim_{n\to\infty} \mathbb{P}(\widehat{p} \neq 1) = 0$. Hence $\widehat{S}_{\widehat{p}} = \widehat{S}_1 + o_{\mathbb{P}}(1)$ and Lemma A.3, which ensures that the retained critical value satisfies $\mathbb{P}\left(\widehat{S}_1 \geq \widehat{z}(\alpha)\right) \to \alpha$, yield that the test (2.7) is asymptotically of level $\alpha$. $\square$

A.2. **Proof of Theorem 2.** The definition (2.2) of $\widehat{p}$ gives, for any $p \in [1, \overline{p}_n]$,

$$\begin{aligned}\widehat{S}_{\widehat{p}} &= \arg\max_{p\in[1,\overline{p}_n]} \left\{\widehat{S}_p - \widehat{R}_0^2 E(p) - \gamma_n \widehat{R}_0^2 V_\Delta(p)\right\} + \widehat{R}_0^2 E(\widehat{p}) + \gamma_n \widehat{R}_0^2 V_\Delta(\widehat{p}) \\ &\geq \widehat{S}_p - \widehat{R}_0^2 E(p) - \gamma_n \widehat{R}_0^2 V_\Delta(p).\end{aligned}$$

Note that this bound implies (3.6). Since the critical value $\widehat{z}(\alpha)$ in (2.7) is bounded under $\mathcal{H}_1$ by Lemma A.3, it is sufficient to find a $p_n \in [1, \overline{p}_n]$ such that $\widehat{S}_{p_n} - \widehat{R}_0^2 E(p_n) - \gamma_n \widehat{R}_0^2 V_\Delta(p_n) \xrightarrow{\mathbb{P}} +\infty$. Let $p_n = 2P_n$ where $P_n$ is as in (3.4). Set

$$\mathcal{R}_n^2 = \sum_{j=1}^{P_n} \left(\frac{R_{j,n}}{R_{0,n}}\right)^2.$$

The detection condition (3.4) gives

$$n\mathcal{R}_n^2 \geq n\rho_n^2 \sum_{j=1}^{P_n} \mathbb{I}\left\{ \left(\frac{R_{j,n}}{R_{0,n}}\right)^2 \geq \rho_n^2 \right\} = nN_n\rho_n^2 \geq \frac{\kappa_*^2 \gamma_n p_n^{1/2}}{2^{1/2}} \to \infty, \tag{A.3}$$

with a constant $\kappa_*$ which can be chosen as large as needed. Lemmas A.2, A.4, Assumption P which ensures $P_n = o\left(n^{1/2}\right)$ and $\gamma_n = o\left(n^{1/4}\right)$, and Proposition A.1 for the case of residuals yield that

$$\begin{aligned}
&\widehat{S}_{p_n} - \widehat{R}_0^2 E\left(p_n\right) - \gamma_n \widehat{R}_0^2 V_\Delta\left(p_n\right) \\
&= \widetilde{S}_{p_n} + O_{\mathbb{P}}\left(1 + n^{1/2} R_{0,n}\mathcal{R}_n\right) - R_{0,n}^2 E\left(p_n\right) - \gamma_n R_{0,n}^2 V_\Delta\left(p_n\right) + O_{\mathbb{P}}\left(\frac{p_n + \gamma_n p_n^{1/2}}{n^{1/2}}\right) \\
&\geq \widetilde{S}_{p_n} + O_{\mathbb{P}}\left(1 + n^{1/2} R_{0,n}\mathcal{R}_n\right) - R_{0,n}^2 E\left(p_n\right) - C\gamma_n R_{0,n}^2 p_n^{1/2}.
\end{aligned}$$

Now the Chebyshev inequality, Propositions A.4 and A.5, give

$$\widetilde{S}_{p_n} = \mathbb{E}\left[\widetilde{S}_{p_n}\right] + O_{\mathbb{P}}\left(\mathrm{Var}^{1/2}\left(\widetilde{S}_{p_n}\right)\right) \geq R_{0,n}^2 E\left(p_n\right) + C' R_{0,n}^2 n\mathcal{R}_n^2 + O_{\mathbb{P}}\left(p_n^{1/2} + n^{1/2}\mathcal{R}_n\right).$$

Hence substituting gives, since $n\mathcal{R}_n^2 \to \infty$ by (A.3),

$$\widehat{S}_{p_n} - \widehat{R}_0^2 E\left(p_n\right) - \gamma_n \widehat{R}_0^2 V_\Delta\left(p_n\right) \geq C' R_{0,n}^2 n\mathcal{R}_n^2\left(1 + o_{\mathbb{P}}(1)\right) - C\gamma_n R_{0,n}^2 p_n^{1/2}\left(1 + o_{\mathbb{P}}(1)\right).$$

Since Assumption R ensures that $R_{0,n}^2$ stays bounded away from 0, (A.3) gives that $\widehat{S}_{p_n} - \widehat{R}_0^2 E\left(p_n\right) - \gamma_n \widehat{R}_0^2 V_\Delta\left(p_n\right) \xrightarrow{\mathbb{P}} +\infty$ as requested provided $\kappa_*^2 > C'/C$. $\square$

A.3. **Proof of Theorem 3.** Consider first the null hypothesis. As seen from the proof of Theorem 1, it suffices to show that

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{p\in[2,\overline{p}_n]} \frac{(\widehat{S}_p^* - \widehat{S}_1^*) - E_\Delta(p)}{V_\Delta(p)} \geq \gamma_n\right) = 0,$$

a statement which implies that $\widehat{p}^{*}=1+o_{\mathbb{P}}(1)$ so that Lemma A.3 implies that the conclusion of Theorem 1 holds for the test based upon $\widehat{S}^{*}_{\widehat{p}^{*}}$. Since $|R_{j,n}|\le\|u_{t,n}\|_2\left\|u_{t,n}-\overline{u}^{t-j}_{t,n}\right\|_2$ and

$$\begin{aligned}\mathbb{E}\left[u^2_{t-j,n}u^2_{t-j,n}\right]&=\mathbb{E}\left[\left(\overline{u}^{t-j}_{t,n}\right)^2u^2_{t-j,n}\right]+\mathbb{E}\left[\left(u^2_{t,n}-\left(\overline{u}^{t-j}_{t,n}\right)^2\right)u^2_{t-j,n}\right]\\&=R^2_{0,n}+\mathbb{E}\left[\left(u_{t,n}-\overline{u}^{t-j}_{t,n}\right)\left(u_{t,n}+\overline{u}^{t-j}_{t,n}\right)u^2_{t-j,n}\right],\end{aligned}$$

(A.2) shows

$$\left|\tau^2_{j,n}-R^2_{0,n}\right|\le C\|u_{t,n}\|_8^3\,\Theta_2(j)\le Cj^{-6}\tag{A.4}$$

for all $j\ge1$. Now Lemmas A.2 and A.4, Assumptions K, P and R, and Proposition A.1 give

$$\begin{aligned}&\max_{p\in[2,\overline{p}_n]}\frac{\left|(\widehat{S}^*_p-\widehat{S}^*_1)-(\widehat{S}_p-\widehat{S}_1)/\widehat{R}^2_0\right|}{V_\Delta(p)}\le C\max_{p\in[1,\overline{p}_n]}\frac{\left|\widehat{S}^*_p-\widehat{S}_p/\widehat{R}^2_0\right|}{p^{1/2}}\\&\le C\max_{p\in[1,\overline{p}_n]}\frac{n}{p^{1/2}}\sum_{j=1}^{p}\left(\frac{\widehat{R}_j}{\widehat{R}_0}\right)^2\left\{\left|\frac{\widehat{\tau}^2_j}{\widehat{R}^2_0}-\frac{\tau^2_{j,n}}{R^2_{0,n}}\right|+\left|\frac{\tau^2_{j,n}}{R^2_{0,n}}-1\right|\right\}\\&\le Cn\overline{p}_n^{1/2}O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{3/2}\right)+O_{\mathbb{P}}(1)\,n\sum_{j=1}^{\overline{p}_n}\frac{\widehat{R}^2_j}{j^6}\\&=o_{\mathbb{P}}(1)+O_{\mathbb{P}}\left(\sum_{j=1}^{\overline{p}_n}\frac{\mathrm{Var}\left(n^{1/2}\widehat{R}_j\right)}{j^6}\right)=O_{\mathbb{P}}(1).\end{aligned}$$

Hence (3.1) and Proposition A.3

$$\begin{aligned}&\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\frac{(\widehat{S}^*_p-\widehat{S}^*_1)-E_\Delta(p)}{V_\Delta(p)}\ge\gamma_n\right)\\&=\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\frac{(\widehat{S}_p-\widehat{S}_1)/\widehat{R}^2_0-E_\Delta(p)}{V_\Delta(p)}+O_{\mathbb{P}}(1)\ge\gamma_n\right)\\&\le\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\frac{(\widehat{S}_p-\widehat{S}_1)/\widehat{R}^2_0-E_\Delta(p)}{V_\Delta(p)}\ge\left(1+\frac{\epsilon}{2}\right)(2\ln\ln n)^{1/2}\right)+o(1)\\&=o(1),\end{aligned}$$

which gives the desired result under $\mathcal{H}_0$.

Consider now Theorem 2 and $\mathcal{H}_1$. Define

$$\widehat{S}_p^{\bigstar} = n \sum_{j=1}^{p} K^2\left(\frac{j}{p}\right)\frac{\widehat{R}_j^2}{\tau_{j,n}^2}, \quad \widetilde{S}_p^{\bigstar} = n \sum_{j=1}^{p} K^2\left(\frac{j}{p}\right)\frac{\widetilde{R}_j^2}{\tau_{j,n}^2}.$$

Let $P_n$ be as in (3.4) and define $p_n = 2P_n$ and $\mathcal{R}_n$ as in the proof of Theorem 2. Then Assumptions K and R, Propositions A.1 and A.2

$$\left|\widehat{S}_{p_n}^{*} - \widehat{S}_{p_n}^{\bigstar}\right| \leq Cn \sum_{j=1}^{p_n} \frac{\widehat{R}_j^2}{\tau_{j,n}^2}\left|\frac{\tau_{j,n}^2}{\widehat{\tau}_j^2} - 1\right| = O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)\check{S}_{p_n}^{\bigstar},$$

$$\left|\widehat{S}_{p_n}^{\bigstar} - \widetilde{S}_{p_n}^{\bigstar}\right| \leq C\left|\widehat{S}_{p_n} - \widetilde{S}_{p_n}\right| = O_{\mathbb{P}}\left(n^{1/2}\mathcal{R}_n\right).$$

Hence, for observed variables or residuals,

$$\widehat{S}_{p_n}^{*} = \left(1 + O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)\right)\widetilde{S}_{p_n}^{\bigstar} + O_{\mathbb{P}}\left(n^{1/2}\mathcal{R}_n\right)$$

The proof now follows the steps of the one of Theorem 2 based on the order above, Proposition A.4 and A.5, and Lemma A.4 which gives $\mathbb{E}\left[\widetilde{S}_{p_n}^{\bigstar}\right] \leq C\left(p_n + n\mathcal{R}_n^2\right)$. Hence, since $p_n = o\left((\log n/n)^{1/2}\right)$,

$$\begin{aligned}
\widehat{S}_{\widehat{p}^*}^{*} &= \arg\max_{p\in[1,\overline{p}_n]}\left\{\widehat{S}_p^{*} - E(p) - \gamma_n V_{\Delta}(p)\right\} + E(\widehat{p}^*) + \gamma_n V_{\Delta}(\widehat{p}^*) \\
&\geq \widehat{S}_{p_n}^{*} - E(p_n) - C\gamma_n p_n^{1/2} \\
&= \left(1 + O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)\right)\left(\mathbb{E}\left[\widetilde{S}_{p_n}^{\bigstar}\right] + \mathrm{Var}^{1/2}\left(\widetilde{S}_{p_n}^{\bigstar}\right)\right) - E(p_n) - C\gamma_n p_n^{1/2} \\
&= C' R_{0,n}^2 n\mathcal{R}_n^2 - C\gamma_n R_{0,n}^2 p_n^{1/2} + O_{\mathbb{P}}\left(p_n^{1/2} + n^{1/2}\mathcal{R}_n + \left(\frac{\log n}{n}\right)^{1/2}\left(p_n + n\mathcal{R}_n^2\right)\right) \\
&= C' R_{0,n}^2 n\mathcal{R}_n^2\left(1 + o_{\mathbb{P}}(1)\right) - C\gamma_n R_{0,n}^2 p_n^{1/2}\left(1 + o_{\mathbb{P}}(1)\right) \xrightarrow{\mathbb{P}} +\infty
\end{aligned}$$

provided $\kappa_*$ is large enough. $\square$

A.4. **Proof of Theorem 4.** Since $\mathbb{P}(\widehat{p}^* = 1) \to 1$ under $\mathcal{H}_0$, condition (A0) and (3.7) give

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\left(\widehat{S}^*_{\widehat{p}^*} \geq \widehat{c}^*_n(\alpha)\right) &= \lim_{n\to\infty} \mathbb{P}\left(\widehat{S}^*_1 \geq \widehat{c}^*_n(\alpha)\right) = \lim_{n\to\infty} \mathbb{P}\left(\widehat{S}^*_1 \geq \widehat{S}^*_1 - \widehat{T}_n + \widehat{t}_n(\alpha)\right) \\ &= \lim_{n\to\infty} \mathbb{P}\left(\widehat{T}_n \geq \widehat{t}_n(\alpha)\right) = \alpha,\end{aligned}$$

so that the test of interest is asymptotically of level $\alpha$. Let us now consider the alternative. Arguing as in the proof of Theorems 2 and 3 under condition (A1) shows that the test with critical value $\widehat{c}_n(\alpha)$ detects the alternatives (3.4) provided $\kappa_*$ is taken large enough. Consider now (3.8). The definition of (2.9) gives, since $E_\Delta(\widehat{p}^*) + \gamma_n V_\Delta(\widehat{p}^*) \geq 0$ under Assumption K,

$$\begin{aligned}\widehat{S}^*_{\widehat{p}^*} &= \max_{p\in[1,\overline{p}_n]} \left(\widehat{S}^*_p - E_\Delta(p) - \gamma_n V_\Delta(p)\right) + E_\Delta(\widehat{p}^*) + \gamma_n V_\Delta(\widehat{p}^*) \\ &\geq \widehat{S}^*_1 - E_\Delta(1) - \gamma_n V_\Delta(1) = \widehat{S}^*_1.\end{aligned}$$

Hence, by (3.7)

$$\mathbb{P}\left(\widehat{S}^*_{\widehat{p}^*} \geq \widehat{c}_n(\alpha)\right) \geq \mathbb{P}\left(\widehat{S}^*_1 \geq \widehat{c}_n(\alpha)\right) = \mathbb{P}\left(\widehat{S}^*_1 \geq \widehat{S}^*_1 - \widehat{T}_n + \widehat{t}_n(\alpha)\right) = \mathbb{P}\left(\widehat{T}_n \geq \widehat{t}_n(\alpha)\right),$$

which is (3.8). □

A.5. **Proof of Theorem 5.** We first introduce a set of alternatives. Let $f(\cdot)$ denote the spectral density of a centered Gaussian stationary process $\{u_t\}$ .with covariance coefficients $R_j$. Define a Hölder class of processes as

$$\text{Hölder}(L) = \left\{\{u_t\} : 1/3 \leq \inf_{\lambda\in[-\pi,\pi]} f(\lambda) \leq \sup_{\lambda\in[-\pi,\pi]} f(\lambda) \leq 3, \ \sup_{\lambda\in[-\pi,\pi]} |f'(\lambda)| \leq L, \ \sum_{j=0}^{\infty} |R_j| \leq L\right\}.$$

The next Lemma describes a family of alternatives which satisfies Assumption R uniformly for prescribed constants and a given $\delta_a(j)$.

**Lemma A.5.** *Consider a centered stationary Gaussian process* $\{u_t\}$ *with spectral density function* $f(\lambda) = \exp(g(\lambda))/(2\pi)$, *where*

$$g(\lambda) = 2\rho \sum_{k=1}^{p} b_k \cos(k\lambda), \qquad b_k = -1, 0, 1. \tag{A.5}$$

*If $p \geq 1$ and $\rho \geq 0$ are such that $p^2\rho \leq \epsilon \leq 1/6$ then there is some constant $L > 0$, independent of $\epsilon$, $p$, $\rho$ and $b = (b_k, k \in [1,p])$, such that (i) $|R_0 - 1| \leq 6\rho\epsilon$ and $|R_j - \rho b_j| \leq 6\rho\epsilon$ for $j \in [1,p]$; (ii) $|R_j| \leq 3\rho\,(2\epsilon)^{\ell}$ for all $j$ in $[\ell p + 1, (\ell+1)\,p)$ and all $\ell \geq 1$; (iii) $\{u_t\}$ is in Hölder$(L)$; (iv) Suppose that $\rho_n^2 = \rho_n^2(p) = 2\kappa_n^2\,(2\log\log n)^{1/2} / \left(np^{1/2}\right)$ for some $\kappa_n > 0$ and bounded away from infinity, and that $p \in [1, P_n]$ with $P_n = o\left(\left(n/\left(\kappa_n^2 \log\log n\right)^{1/2}\right)^{1/14}\right)$. Then the associated family of processes $\{u_t\,(b,p)\,; b \in \{-1,0,1\}^p, p \in [1,P_n]\}$ satisfies Assumption R for any $a > 0$ and a $\delta_a\,(j) = O\left(j^{-7-1/4}\right)$.*

**Proof of Lemma A.5.** Rewrite $g$ as $g\,(\lambda) = \rho\sum_{k=-p}^{p} b_k \exp\,(ik\lambda)$, $b_0 = 0$, $b_k = b_{-k} = b_{|k|}$. Since $\exp\,(x) = \sum_{m=0}^{\infty} x^m/m!$ uniformly over any compact set and $\max_\lambda |g\,(\lambda)| \leq 2p\rho \leq 2\epsilon \leq 1/3$, we have

$$R_j = \int_{-\pi}^{\pi} \exp\,(-ij\lambda)\, f\,(\lambda)\, d\lambda = \frac{1}{2\pi}\sum_{m=0}^{\infty}\frac{1}{m!}\int_{-\pi}^{\pi}\exp\,(-ij\lambda)\,(g\,(\lambda))^m\, d\lambda. \tag{A.6}$$

For $m > 0$, since $\int_{-\pi}^{\pi}\exp\,(-ij\lambda)\,d\lambda = 2\pi$ if $j = 0$ and $0$ if $j \neq 0$,

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\,(-ij\lambda)\,(g\,(\lambda))^m\,d\lambda \\
&= \frac{\rho^m}{2\pi}\sum_{(k_1,\ldots,k_m)\in K_m} b_{k_1}\times\cdots\times b_{k_m}\int_{-\pi}^{\pi}\exp\,(i\,(k_1+\ldots+k_m-j)\,\lambda)\,d\lambda \\
&= \rho^m \sum_{(k_1,\ldots,k_m)\in K_m(j)} b_{k_1}\times\cdots\times b_{k_m},
\end{aligned} \tag{A.7}$$

where $K_m$ is the set of $m$-tuples with entries in $[-p,p]\setminus\{0\}$ so that $\#K_m = (2p)^m$ and $K_m\,(j)$ contains $m$-tuples in $K_m$ for which $k_1+\cdots+k_m = j$ so that $\#K_m(j) \leq (2p)^{m-1}$.

*Proof of (i).* Part (i) is a consequence of (A.6), (A.7) and inequality $2p\rho \leq 2\epsilon < 1$ which together imply that for $j \in [0,p]$, $|R_j - \mathbb{I}\,(j=0) - \rho b_j| \leq \rho\sum_{m=2}^{\infty}\frac{(2p\rho)^{m-1}}{m!} \leq 2p\rho^2\sum_{m=0}^{\infty}\frac{1}{m!} \leq 2e\rho\epsilon < 6\rho\epsilon$.

*Proof of (ii).* Let $\ell p + 1 \leq j > (\ell+1)\,p$. Observe that $K_m\,(j)$ is an empty set when $m \leq \ell$. Hence it follows from (A.6) and (A.7) that $|R_j| \leq \left|\frac{1}{2\pi}\sum_{m=\ell+1}^{\infty}\frac{1}{m!}\int_{-\pi}^{\pi}\exp\,(-ij\lambda)\,(g\,(\lambda))^m\,d\lambda\right| \leq \rho\sum_{m=\ell+1}^{\infty}\frac{(2p\rho)^{m-1}}{m!} \leq \rho\,(2\epsilon)^{\ell}\,e$.

*Proof of (iii).* Observe that $|g(\lambda)| \leq 2\rho p \leq 2\epsilon \leq 1/3$ and that therefore

$$1/3 < 1 - 1/3 < \exp(-1/3) \leq f(\lambda) \leq \exp(1/3) \leq e \leq 3 \qquad \text{for all } \lambda \in [-\pi, \pi].$$

Parts (i), (ii) and $0 \leq \rho \leq \epsilon < 1/6$, $p\rho \leq 1/6$ yield that, for $L$ large enough,

$$\begin{aligned}\sum_{j=0}^{\infty} |R_j| &\leq R_0 + \sum_{j=1}^{p} |R_j| + \sum_{\ell=1}^{\infty} \sum_{j=\ell p+1}^{(\ell+1)p} |R_j| \leq 1 + 6\rho\epsilon + (1+6\epsilon)\, p\rho + 3 \sum_{\ell=1}^{\infty} (\ell+1)\, p\rho\, (2\epsilon)^{\ell} \\ &\leq 1 + 1 + 1 + 1 + \sum_{\ell=1}^{\infty} (\ell+1)(2\epsilon)^{\ell} \leq L.\end{aligned}$$

Since $f'(\lambda) = g'(\lambda) f(\lambda)$ with $g'(\lambda) = -2\rho \sum_{k=1}^{p} b_k k \sin(k\lambda)$, we have $\sup_{\lambda \in [-\pi,\pi]} |f'(\lambda)| \leq 3 \times 2p^2\rho \leq 1$.

*Proof of (iv).* Let $u_t = \varepsilon_t + \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j}$ be the Wold decomposition of the process. Brillinger (2001) and $\int_{-\pi}^{\pi} \log f(\lambda) \exp(ij\lambda)\, d\lambda / 2\pi = \rho b_j$ gives

$$\psi_j = \frac{\int_{-\pi}^{\pi} \exp\left(\rho \sum_{k=1}^{p} b_k \exp(-ik\lambda)\right) \exp(ij\lambda)\, d\lambda}{\int_{-\pi}^{\pi} \exp\left(\rho \sum_{k=1}^{p} b_k \exp(-ik\lambda)\right) d\lambda},$$

$$\operatorname{Var}(\varepsilon_t) = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left( \rho \sum_{k=1}^{p} b_k \exp(-ik\lambda) \right) d\lambda \right|^2.$$

Arguing as in (i) and (ii) with an expansion as in (A.6) give $\operatorname{Var}(\varepsilon_t) = 1$, $|\psi_j - \rho b_j| \leq C\rho\epsilon$ for $j \in [1, p]$ and $|\psi_j| \leq C\rho(2\epsilon)^{\ell}$ for all $j \in [\ell p + 1, (\ell+1)p)$ and all $\ell \geq 1$. Gaussianity, the choice of $\rho$ in (iv) with the restriction on $P_n$ and Wu (2005) give, for any $a > 1$, $\delta_{12a}(j) \leq C_a |\psi_j| \leq C_a j^{-7-1/4}$. That the other conditions of Assumption R hold uniformly in $p \in [1, P_n]$ follows from (i) and (ii). □

We will now define a family $\mathcal{F}_n$ of correlated Gaussian alternatives. We first introduce some notation. Consider $\widetilde{\gamma}_n = (2 \ln\ln n)^{1/2}$ and $\mathcal{P}' = \{2^j, j = 1, \ldots, J_n\}$, $2^{J_n} = P_n = o\left(\overline{p}_n \wedge (n/\widetilde{\gamma}_n)^{1/14}\right)$ so that $\mathcal{P}' \subset [1, \overline{p}_n]$ for $n$ large enough. Define also

$$\rho_n^2(p) = 2\frac{\kappa_n^2 \widetilde{\gamma}_n}{np^{1/2}}, \quad \widetilde{\rho}_n(p) = 2\rho_n^2(p) \quad \epsilon_n = P_n^2 \rho_n(P_n) = \frac{(\widetilde{\gamma}_n)^{1/2} \kappa_n P_n^{7/4}}{n^{1/2}} = o(1). \tag{A.8}$$

Since $p^2\rho_n(p) \leq \epsilon_n$ for all $p \in \mathcal{P}'$, $\epsilon_n$ plays the role of the real number $\epsilon$ of Lemma A.5 and we assume from now on that $n$ is so large that $\epsilon_n \leq 1/6$. Consider the following log-spectral density functions:

$$g(\lambda; b, p) = 2\widetilde{\rho}_n(p) \sum_{k \in [p,2p)} b_k \cos(k\lambda), \quad b = (b_1, \ldots, b_{P_n}) \in \{-1, 1\}^{P_n}, \quad p \in \mathcal{P}'.$$

Functions $g$ are of the form specified in (A.5). Let $W$ be a symmetric standard Brownian motion process. Consider a centered stationary Gaussian processes

$$u_{t,n}(b, p) = \frac{1}{(2\pi)^{1/2}} \int_{-\pi}^{\pi} \exp\left(\frac{g(\lambda; b, p)}{2}\right) \exp(it\lambda)\, dW(\lambda).$$

Observe that $u_{t,n}(0, p)$ does not depend on $p$ and is a Gaussian white noise process with variance 1. Let $\{R_{j,n}(b, p)\}$ denote the covariance function of $u_{t,n}(b, p)$. The family $\mathcal{F}_n$ of Gaussian processes can now be defined as

$$\mathcal{F}_n = \left\{ \{u_{t,n}(b, p)\}, b \in \{-1, 1\}^{P_n}, p \in \mathcal{P}' \right\}.$$

Lemma A.5 implies that all sequences $\{u_{t,n}\}$ in $\mathcal{F}_n$ satisfies Assumption R and that $\mathcal{F}_n \subset$Hölder$(L)$. We now study the asymptotic behavior of the stochastic covariance sequence $\{R_{j,n}(B, P)\}$. Let $N_n(b, p)$ be as in (3.3), that is

$$N_n(b, p) = N_n(\{u_{t,n}(b, p)\}, p, \rho_n(p)) = \# \left\{ \left| \frac{R_{j,n}(b, p)}{R_{0,n}(b, p)} \right| \geq \rho_n(p),\ j \in [1, p] \right\}.$$

Lemma A.5-(i,ii) and (A.8) gives that $N_n(b, p) = p/2$ for $n$ large enough and uniformly in $p = 2^j \in \mathcal{P}'$, so that $\rho_n^2(p) = 2\kappa_n^2 \widetilde{\gamma}_n / \left(np^{1/2}\right) = \kappa_n^2 \widetilde{\gamma}_n p^{1/2} / (nN_n(b, p))$. Hence the sequences $\{u_{t,n}\}$ in $\mathcal{F}_n$ satisfies condition (i) in Theorem 5. Therefore the Theorem will be proved if we show that $\sup_{T_n} \min_{\{u_{t,n}\} \in \mathcal{F}_n} \mathbb{P}(T_n = 0) \leq \alpha + o(1)$, where $\sup_{T_n}$ is a supremum over asymptotically $\alpha$-level tests. Since the equivalence result of Golubev et al. (2010) holds over $\mathcal{F}_n \subset$Hölder$(L)$ this is equivalent to show that $\sup_{T_n} \min_{\{U_n\} \in \mathcal{F}_n} \mathbb{Q}(T_n = 0) \leq \alpha + o(1)$, $\mathbb{Q}$

being the distribution of the continuous time regression model

$$dU_n(\lambda; b, p) = g(\lambda; b, p)\, d\lambda + 2\pi^{1/2} \frac{dW(\lambda)}{n^{1/2}}, \qquad \lambda \in [-\pi, \pi],$$

where $W(\cdot)$ is a Brownian motion over $\lambda \in [-\pi, \pi]$. This can be done as in Spokoiny (1996, Proof of Theorem 2.3) by bounding $\sup_{T_n} \min_{\{U_n\} \in \mathcal{F}_n} \mathbb{Q}(T_n = 0)$ with a Bayes risk, based on the choice of a uniform distribution for $p$ and a Bernoulli one for $b$. $\square$

A.6. **Proof of Lemma 1.** The first approximation $R_{0,n} = \sigma^2 \left(1 + O\left(\gamma_n P_n^{1/2}/n\right)\right)$ follows easily from the definition (4.1) of the alternative. To show that the second approximation is valid, note that for $j = 1, ..., P_n$,

$$R_{j,n} = \frac{\nu \gamma_n^{1/2}}{n^{1/2} P_n^{1/4}} \psi_j \sigma^2 + \left( \frac{\nu \gamma_n^{1/2}}{n^{1/2} P_n^{1/4}} \right)^2 (\psi_{j+1}\psi_1 + \cdots + \psi_{P_n}\psi_{P_n - j})\, \sigma^2.$$

By the Cauchy-Schwarz inequality, $|\psi_{j+1}\psi_1 + \cdots + \psi_{P_n}\psi_{P_n-j}| \leq \sum_{k=1}^{P_n} \psi_k^2 = O(P_n)$ for all $j = 1, ..., P_n$, hence, uniformly in $j = 1, ..., P_n$,

$$R_{j,n} = \frac{\nu \gamma_n^{1/2}}{n^{1/2} P_n^{1/4}} \psi_j \sigma^2 + O\left( \frac{\gamma_n P_n^{1/2}}{n} \right) = \frac{\nu \gamma_n^{1/2}}{n^{1/2} P_n^{1/4}} \psi_j \sigma^2 + o\left( \frac{\gamma_n^{1/2}}{n^{1/2} P_n^{1/4}} \right)$$

since $P_n = o((n/\gamma_n)^{2/3})$.

For the expression of $\mathbb{E}[u_t | u_{t-k}, k \geq 1]$, observe that (4.1) gives, for $n$ large enough,

$$\begin{aligned}
\mathbb{E}[u_t | u_{t-k}, k \geq 1] &= \frac{\nu \gamma_n^{1/2}}{n^{1/2} P_n^{1/4}} \sum_{k=1}^{P_n} \psi_k \varepsilon_{t-k} \\
&= \frac{\nu \gamma_n^{1/2}}{n^{1/2} P_n^{1/4}} \sum_{k=1}^{P_n} \psi_k \left( u_{t-k} - \frac{\nu \gamma_n^{1/2}}{n^{1/2} P_n^{1/4}} \sum_{j=1}^{P_n} \psi_j \varepsilon_{t-k-j} \right) \\
&= \frac{\nu \gamma_n^{1/2}}{n^{1/2} P_n^{1/4}} \sum_{k=1}^{P_n} \psi_k u_{t-k} - \frac{\nu^2 \gamma_n}{n P_n^{1/2}} \sum_{k=1}^{P_n} \psi_k \sum_{j=1}^{P_n} \psi_j \varepsilon_{t-k-j}.
\end{aligned}$$

Now, since $\{\varepsilon_t\}$ is a strong white noise and $\sum_{k=1}^{P_n}\psi_k^2 = O(P_n)$,

$$\frac{\nu^2\gamma_n}{nP_n^{1/2}}\sum_{k=1}^{P_n}\psi_k\sum_{j=1}^{P_n}\psi_j\varepsilon_{t-k-j} = \frac{\nu^2\gamma_n}{nP_n^{1/2}}\sum_{\ell=2}^{2P_n}\left(\sum_{k=1}^{\max(P_n,\ell-1)}\psi_k\psi_{\ell-k}\right)\varepsilon_{t-\ell}$$

$$= O_{\mathbb{P}}\left(\left(\frac{\gamma_n^2}{n^2P_n}\sum_{\ell=2}^{2P_n}\left(\sum_{k=1}^{\max(P_n,\ell-1)}\psi_k\psi_{\ell-k}\right)^2\right)^{1/2}\right)$$

$$= O_{\mathbb{P}}\left(\left(\frac{\gamma_n^2\left(\sum_{k=1}^{P_n}\psi_k^2\right)^2}{n^2}\right)^{1/2}\right) = O_{\mathbb{P}}\left(\frac{\gamma_nP_n}{n}\right),$$

which ends the proof of the Lemma. □

A.7. **Proof of Proposition 1.** Let us now check consistency of the test (2.7) under the assumption that $\min_{k\in[1,P_n]}|\psi_k\sigma^2| \geq 1$. Define $\rho_n = (\nu/2)\gamma_n^{1/2}/\left(n^{1/2}P_n^{1/4}\right)$. Lemma 1 implies that $N_n = P_n(1+o(1))$ for such a $\rho_n$, which therefore satisfies

$$\rho_n = (1+o(1))(\nu/2)\left(\gamma_nP_n^{1/2}/N_n\right)^{1/2}/n^{1/2},$$

so that (3.4) asymptotically holds provided $\nu \geq 3\kappa^*$ and the test is consistent if $1 \leq P_n \leq \overline{p}_n/2$ by Theorem 2 provided the considered alternatives satisfies Assumption R. Wu (2005) gives that the alternative (4.1) satisfies for any $a > 0$,

$$\delta_{12a}(j) \leq C_a\frac{\nu\gamma_n^{1/2}}{n^{1/2}P_n^{1/4}}|\sigma\psi_j| \text{ for all } j\in[1,P_n],\ \delta_{12a}(j) = 0 \text{ for all } j > P_n.$$

Hence the condition $P_n = O\left((n/\gamma_n)^{1/14}\right)$ gives that $\delta_{12a}(j) \leq Cj^{-7-1/4}$ since the $|\sigma\psi_j|$ are bounded away from infinity. Moreover Gaussianity ensures that

$$\|u_{t,n}-\varepsilon_t\|_{12a} \leq C_a\sigma\left(\frac{\nu^2\gamma_n}{nP_n^{1/2}}\sum_{k=1}^{P_n}\psi_k^2\right)^{1/2} = O\left(\frac{\nu\gamma_n^{1/2}P_n^{1/4}}{n^{1/2}}\right) = o(1),$$

which gives $\operatorname{Var}(u_{t,n}) = \sigma^2 + o(1)$ and $\max_{j\in[1,n]}\operatorname{Var}^2(u_{t,n})/\operatorname{Var}(u_{t,n}u_{t+j,n}) = 1 + o(1)$ so that Assumption R holds. This ends the proof of Proposition 1-(i).

Consider now the other tests in Proposition 1-(ii). Define $\widetilde{R}_{1,j} = \sum_{t=1}^{n-j} u_{t,n}u_{t+j,n}/n$, $\widetilde{R}_{0,j} = \sum_{t=1}^{n-j} \varepsilon_t\varepsilon_{t+j}/n$, $\widetilde{\tau}_{1,j}^2 = \sum_{t=1}^{n-j} u_{t,n}^2 u_{t+j,n}^2/(n-j) - n\widetilde{R}_{1,j}^2/(n-j)$ and $\widetilde{\tau}_{0,j}^2 = \sum_{t=1}^{n-j} \varepsilon_t^2\varepsilon_{t+j}^2/(n-j) - n\widetilde{R}_{0,j}^2/(n-j)$. Define also $\eta_t = \eta_{t,n} = \nu\sum_{k=1}^{\infty}\psi_k\varepsilon_{t-k}$, setting $\psi_k = 0$ for $k > P_n$, so that $u_{t,n} = \varepsilon_t + \gamma_n^{1/2}\eta_t/\left(n^{1/2}P_n^{1/4}\right)$. We have

$$\left|\widetilde{R}_j - \widetilde{R}_{0,j}\right| \leq \frac{\gamma_n^{1/2}}{n^{3/2}P_n^{1/4}}\left|\sum_{t=1}^{n-j}\eta_t\varepsilon_{t+j}\right| + \frac{\gamma_n^{1/2}}{n^{3/2}P_n^{1/4}}\left|\sum_{t=1}^{n-j}\varepsilon_t\eta_{t+j}\right| + \frac{\gamma_n}{n^2P_n^{1/2}}\left|\sum_{t=1}^{n-j}\eta_t\eta_{t+j}\right|.$$

The Burkholder inequality gives, for any $a > 1$,

$$\left\|\frac{\gamma_n^{1/2}}{n^{3/2}P_n^{1/4}}\sum_{t=1}^{n-j}\eta_t\varepsilon_{t+j}\right\|_a \leq C\frac{\gamma_n^{1/2}(n-j)^{1/2}}{n^{3/2}P_n^{1/4}}\|\eta_t\|_a \leq C\frac{\gamma_n^{1/2}P_n^{1/4}}{n},$$

$$\left\|\frac{\gamma_n^{1/2}}{n^{3/2}P_n^{1/4}}\sum_{t=1}^{n-j}\left(\varepsilon_t\eta_{t+j} - \psi_j\varepsilon_t^2\right)\right\|_a \leq \left\|\frac{\gamma_n^{1/2}}{n^{3/2}P_n^{1/4}}\sum_{t=1}^{n-j}\varepsilon_t\left(\sum_{k=0}^{j-1}\psi_j\varepsilon_{t+j-k}\right)\right\|_a$$

$$+\left\|\frac{\gamma_n^{1/2}}{n^{3/2}P_n^{1/4}}\sum_{t=1}^{n-j}\left(\sum_{k=j+1}^{\infty}\psi_j\varepsilon_{t+j-k}\right)\varepsilon_t\right\|_a \leq C\frac{\gamma_n^{1/2}P_n^{1/4}}{n},$$

$$\left\|\frac{\gamma_n^{1/2}}{n^{3/2}P_n^{1/4}}\sum_{t=1}^{n-j}\left(\varepsilon_t^2 - \sigma^2\right)\right\|_a \leq C\frac{\gamma_n^{1/2}}{nP_n^{1/4}}, \quad \left\|\frac{\gamma_n}{n^2P_n^{1/2}}\sum_{t=1}^{n}\eta_t^2\right\|_a \leq \frac{\gamma_n}{nP_n^{1/2}} \leq C\frac{\gamma_nP_n^{1/2}}{n},$$

for all $j$. Note also that $\left|\sum_{t=1}^{n-j}\eta_t\eta_{t+j}\right| \leq \sum_{t=1}^{n}\eta_t^2$ and the Markov inequality give for $a$ large enough, since $\gamma_nP_n^{1/2} = o(n^{1/4})$

$$\max_{j\in[1,n]}\left|\widetilde{R}_{1,j} - \widetilde{R}_{0,j}\right|^a = O_{\mathbb{P}}\left(\max_{j\in[1,n]}\left|\widetilde{R}_{1,j} - \widetilde{R}_{0,j}\right|^a\right)$$

$$= O_{\mathbb{P}}\left(\sum_{j=1}^{n}\left\|\frac{\gamma_n^{1/2}}{n^{3/2}P_n^{1/4}}\sum_{t=1}^{n-j}\eta_t\varepsilon_{t+j} + \sum_{t=1}^{n-j}\varepsilon_t\eta_{t+j}\right\|_a^a + \left\|\frac{\gamma_n}{n^2P_n^{1/2}}\sum_{t=1}^{n}\eta_t^2\right\|_a^a\right)$$

$$= O_{\mathbb{P}}\left(n\left(\frac{\gamma_n^{1/2}P_n^{1/4}}{n}\right)^a + \left(\frac{\gamma_nP_n^{1/2}}{n}\right)^a\right) = o_{\mathbb{P}}\left(\frac{1}{n^{7a/8-1}} + \frac{1}{n^{3a/4}}\right)$$

$$= o_{\mathbb{P}}\left(\frac{1}{(n\log n)^{a/2}}\right).$$

Hence

$$\max_{j\in[1,n]}\left|\widetilde{R}_{1,j}-\widetilde{R}_{0,j}\right| = o_{\mathbb{P}}\left(\frac{1}{(n\log n)^{1/2}}\right). \tag{A.9}$$

Arguing similarly for the $\widetilde{\tau}_{k,j}^2$ give, since $J_n = O\left(n^{1/2}\right)$

$$\max_{j\in[1,J_n]}\left|\widetilde{\tau}_{1,j}^2-\widetilde{\tau}_{0,j}^2\right| = o_{\mathbb{P}}\left(\frac{1}{(n\log n)^{1/2}}\right) \quad , \max_{j\in[1,J_n]}\left|\widetilde{\tau}_{0,j}^2-\sigma^4\right| = O_{\mathbb{P}}\left(\frac{\log^{1/2} n}{n^{1/2}}\right), \tag{A.10}$$

where the latter is from Proposition A.1. Note that (A.9) and (A.10) gives (4.5). Let $W_{k,n}$, $CvM_{k,n}$, $EL_{k,n}$ be the statistic computed under $G_k$, $k=0,1$, i.e. with $\widetilde{R}_{0,j}/\widetilde{\tau}_{0,j}$ and $\widetilde{R}_{1,j}/\widetilde{\tau}_{1,j}$. Note that (A.9) and (A.10) gives $W_{1,n} = W_{0,n} + o_{\mathbb{P}}(1)$. (4.5) and Proposition A.1 give

$$\begin{aligned}
|CvM_{1,n} - CvM_{0,n}| &\leq \frac{2}{\pi^2}\sum_{j=1}^{J_n}\frac{n\left|\left(\widetilde{R}_{1,j}/\widetilde{\tau}_{1,j}+\widetilde{R}_{0,j}/\widetilde{\tau}_{0,j}\right)\left(\widetilde{R}_{1,j}/\widetilde{\tau}_{1,j}-\widetilde{R}_{0,j}/\widetilde{\tau}_{0,j}\right)\right|}{j^2}\\
&\leq 2\max_{j\in[1,J_n]}\frac{\left|n^{1/2}\widetilde{R}_{0,j}\right|}{\widetilde{\tau}_{0,j}}\times\max_{j\in[1,J_n]}\left|n^{1/2}\left(\frac{\widetilde{R}_{1,j}}{\widetilde{\tau}_{1,j}}-\frac{\widetilde{R}_{0,j}}{\widetilde{\tau}_{0,j}}\right)\right|\frac{2}{\pi^2}\sum_{j=1}^{J_n}\frac{1}{j^2}\\
&+\max_{j\in[1,J_n]} n\left(\frac{\widetilde{R}_{1,j}}{\widetilde{\tau}_{1,j}}-\frac{\widetilde{R}_{0,j}}{\widetilde{\tau}_{0,j}}\right)^2\frac{2}{\pi^2}\sum_{j=1}^{J_n}\frac{1}{j^2}\\
&= n^{1/2}O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)n^{1/2}o_{\mathbb{P}}\left(\frac{1}{(n\log n)^{1/2}}\right)+no_{\mathbb{P}}\left(\frac{1}{n\log n}\right) = o_{\mathbb{P}}(1),
\end{aligned}$$

Hence $CvM_{1,n} = CvM_{0,n} + o_{\mathbb{P}}(1)$. For $EL_n$, $W_{1,n} = W_{0,n} + o_{\mathbb{P}}(1)$ and Xiao and Wu (2011) gives that $\max_{j\in[1,J_n]}\left|\widetilde{R}_{k,j}/\widetilde{\tau}_{k,j}\right| \leq (2\ln n)^{1/2}(1+o_{\mathbb{P}}(1))$ for $k=0,1$ so that $\mathbb{P}\left(\widehat{\gamma}_{EL}^* = \ln n\right) \to 1$ under $G_0$ and $G_1$.We now show that $\mathbb{P}\left(\widehat{p}_{EL}^* = 1\right) \to 1$ under $G_0$. Propositions A.4 and A.5, (A.10) give

$$\mathbb{P}\left(\widetilde{p}^{*}_{0,EL}\neq 1\right)=\mathbb{P}\left(\max_{p\in[2,J_n]}\frac{\widetilde{BP}^{*}_{0,p}-\widetilde{BP}^{*}_{0,1}}{p-1}>\ln n\right)+o(1)$$

$$=\mathbb{P}\left((1+o_{\mathbb{P}}(1))\max_{p\in[2,J_n]}\frac{n\sum_{j=2}^{p}\widetilde{R}^2_{0,j}/\sigma^4}{p-1}>\ln n\right)+o(1)$$

$$=\mathbb{P}\left(\frac{n\sum_{j=2}^{p}\widetilde{R}^2_{0,j}/\sigma^4}{p-1}>\frac{1}{2}\ln n\text{ for some }p\in[2,J_n]\right)+o(1)$$

$$\leq\sum_{p=2}^{J_n}\mathbb{P}\left(\frac{n\sum_{j=2}^{p}\left(\widetilde{R}^2_{0,j}/\sigma^4-\mathbb{E}\left[\widetilde{R}^2_{0,j}/\sigma^4\right]\right)}{p-1}>\frac{1}{2}\ln n-\frac{n\sum_{j=2}^{p}\mathbb{E}\left[\widetilde{R}^2_{0,j}/\sigma^4\right]}{p-1}\right)+o(1)$$

$$\leq\sum_{p=2}^{J_n}\frac{\operatorname{Var}\left(\frac{n\sum_{j=2}^{p}\left(\widetilde{R}^2_{0,j}/\sigma^4-\mathbb{E}\left[\widetilde{R}^2_{0,j}/\sigma^4\right]\right)}{p-1}\right)}{\left(\frac{1}{2}\ln n-\frac{1}{p-1}\sum_{j=2}^{p}(1-j/n)\right)^2}+o(1)$$

$$\leq\frac{C}{\log^2 n}\sum_{p=2}^{J_n}\frac{1}{p-1}+o(1)=O\left(\frac{1}{\log n}\right)+o(1)=o(1).$$

Now, observe that Proposition A.1 and (4.5) give

$$\max_{p\in[2,J_n]}\left|\frac{\widetilde{BP}^{*}_{0,p}-\widetilde{BP}^{*}_{0,1}}{p-1}-\frac{\widetilde{BP}^{*}_{1,p}-\widetilde{BP}^{*}_{1,1}}{p-1}\right|\leq\max_{p\in[2,J_n]}\left|\frac{n\sum_{j=2}^{p}\left(\widetilde{R}^2_{0,j}/\widetilde{\tau}^2_{0,j}-\widetilde{R}^2_{1,j}/\widetilde{\tau}^2_{1,j}\right)}{p-1}\right|$$

$$\leq 2\max_{p\in[2,J_n]}\left|n^{1/2}\frac{\widetilde{R}_{0,j}}{\widetilde{\tau}_{0,j}}\right|\times\max_{p\in[2,J_n]}\left|n^{1/2}\left(\frac{\widetilde{R}_{0,j}}{\widetilde{\tau}_{0,j}}-\frac{\widetilde{R}_{1,j}}{\widetilde{\tau}_{1,j}}\right)\right|+\left(\max_{p\in[2,J_n]}\left|n^{1/2}\left(\frac{\widetilde{R}_{0,j}}{\widetilde{\tau}_{0,j}}-\frac{\widetilde{R}_{1,j}}{\widetilde{\tau}_{1,j}}\right)\right|\right)^2$$

$$=n^{1/2}O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)n^{1/2}o_{\mathbb{P}}\left(\frac{1}{(n\log n)^{1/2}}\right)+no_{\mathbb{P}}\left(\frac{1}{n\log n}\right)=o_{\mathbb{P}}(1).$$

This, since arguing as in the bound above gives $\max_{p\in[2,J_n]}\left|\left(\widetilde{BP}^{*}_{0,p}-\widetilde{BP}^{*}_{0,1}\right)/(p-1)\right|=O_{\mathbb{P}}\left(\log^{1/2}n\right)$, implies that $\max_{p\in[2,J_n]}\left|\left(\widetilde{BP}^{*}_{1,p}-\widetilde{BP}^{*}_{1,1}\right)/(p-1)\right|\leq\log n$ with a probability tending to 1 and then $\mathbb{P}\left(\widehat{p}^{*}_{EL}=1\right)\to 1$ under $G_1$. Hence (4.5) gives that $EL_{1,n}=\widetilde{BP}^{*}_{1,1}+o_{\mathbb{P}}(1)=\widetilde{BP}^{*}_{0,1}+o_{\mathbb{P}}(1)=EL_{0,n}+o_{\mathbb{P}}(1)$, so that $EL_n$ converges in distribution to a Chi square one with one degree of freedom under $G_0$ and $G_1$. $\square$

SUPPLEMENTARY MATERIAL B: PROOFS OF INTERMEDIARY RESULTS

The proofs also use the notion of cumulants, see for example Brillinger (2001, p. 19) or Xiao and Wu (2011) for a definition. Let

$$\mathrm{Cum}\left(u_{t_{1,n}}, \ldots, u_{t_{q,n}}\right) = \Gamma_n(t_1, \ldots, t_q)$$

stands for the $q$th cumulants of $\{u_{t,n}\}$. The next theorem on cumulant summability is Theorem 21 in Xiao and Wu (2011). These authors do not formally consider sequences $\{u_{t,n}\}$ but the following result is a straightforward extension of Xiao and Wu (2011).

**Theorem B.1** (Xiao and Wu (2011)). *Suppose $\{u_{t,n}\}$ is stationary for each $n$, with*

$$\sup_n \|u_{t,n}\|_{q+1} < \infty \text{ and } \sup_n \left\|u_{t,n} - u_{t,n}^{t-j}\right\|_q \leq \delta_q(j) \text{ where } \sum_{j=0}^{\infty} j^{q-2}\delta_q(j) < \infty.$$

*Then there is a $\mathcal{C}$ which only depends on $\sup_n \|u_{t,n}\|_{q+1}$ and $\sum_{j=0}^{\infty} j^{q-2}\delta_q(j)$ such that*

$$\sum_{t_2,\ldots,t_q=-\infty}^{\infty} |\Gamma_n(0, t_2, \ldots, t_q)| \leq \mathcal{C}.$$

In what follows, we drop subscript $n$ in expressions like $u_{t,n}$, $R_{j,n}$, $\Gamma_n(\cdot)$ and $\theta_n$ when there is no ambiguity. We denote

$$K_{jp} = K^2\left(\frac{j}{p}\right) - K^2(j) \qquad \text{and} \qquad K_{1n}(p) = \sum_{j=1}^{n-1} K_{jp}. \tag{B.1}$$

B.1. **Proof of Lemma A.2.** (i) The first three bounds of the lemma follow directly from Assumption K which implies that $K^2(j/p) \geq K^2(j)$ for all $j$ and $\mathbb{I}(x \in [0,1/2])/C \leq K^{2q}(x) \leq C\mathbb{I}(x \in [0,1])$ for some $C > 0$. The Cauchy-Schwarz inequality implies that for any $p \in [1, n/2]$, $E_\Delta(p) = \sum_{j=1}^{n-1}\left(1 - \frac{j}{n}\right) K_{jp} \leq K_{1n}(p) \leq p^{1/2}\left(\sum_{j=1}^{n-1} k_j^2(p)\right)^{1/2} \leq Cp^{1/2}V_\Delta(p)$, which is the last bound in (i). (ii) Write $p = 1 + \nu$. Since $p \leq \overline{p}_n \leq n/2$, the support of $K(\cdot)$

is $[0,1]$ and $K(\cdot)$ is a decreasing function, we have

$$V_{\Delta}^{2}(p) \geq \frac{1}{2} \times 2 \sum_{j=2}^{p} K^{2}\left(\frac{j}{p}\right) \geq \sum_{j=1}^{\nu} K^{2}\left(\frac{1+j}{1+\nu}\right) \geq \sum_{j=1}^{\nu} \int_{j}^{j+1} K^{2}\left(\frac{1+x}{1+\nu}\right) dx$$
$$= \int_{1}^{\nu+1} K^{2}\left(\frac{1+x}{1+\nu}\right) dx = \nu \int_{0}^{1} K^{2}\left(\frac{2+z\nu}{1+\nu}\right) dz.$$

The map $\nu \longmapsto (2+z\nu)/(1+\nu)$, $z \in [0,1)$, is decreasing. Hence, for $\nu \geq 2$, $V_{\Delta}^{2}(p) \geq \nu \int_{0}^{1/2} K^{2}\left(\frac{2+2z}{3}\right) dz \geq C(p-1)$. Now $V_{\Delta}^{2}(2) \geq 2\left(K^{2}\left(\frac{1}{2}\right) - K^{2}(1)\right)^{2} > 0$ gives the desired result for $V_{\Delta}(p)$. Since $K$ is nonincreasing, $p \longmapsto E_{\Delta}(p)$ is non decreasing and $E_{\Delta}(p) \geq 0$ for all $p \in \mathcal{P}$. □

B.2. **Proof of Lemma A.3.** Under $\mathcal{H}_0$, The proof repeats the steps of Lee (2007), Lobato (2001) and Kuan and Lee (2006) using the joint FCLT of Assumption M. The joint FCLT of Assumption M gives that the critical values are $O_{\mathbb{P}}(1)$ under $\mathcal{H}_1$. □

B.3. **Proof of Lemma A.4.** Equation (5.3.21) in Priestley (1981) and Theorem B.1 gives uniformly in $j$,

$$\operatorname{Var}\left(\widetilde{R}_{j}\right) = \frac{1}{n} \sum_{j_1=-n+j+1}^{n-j-1} \left(1 - \frac{|j_1|+j}{n}\right)\left(R_{j_1}^{2} + R_{j_1+j}R_{j_1-j} + \Gamma(0, j_1, j, j_1+j)\right)$$
$$\leq \frac{2}{n} \sum_{j_1=-2n}^{2n} R_{j_1}^{2} + \frac{1}{n} \sum_{j_2,j_3,j_4=-\infty}^{+\infty} |\Gamma(0, j_2, j_3, j_4)|$$
$$\leq \frac{4}{n} \sum_{j=0}^{\infty} R_{j}^{2} + \frac{1}{n} \sum_{j_2,j_3,j_4=-\infty}^{+\infty} |\Gamma(0, j_2, j_3, j_4)| < C. \square$$

B.4. **Proof of Proposition A.1.** For the sake of brevity we assume that $\theta$ is unidimensional. That

$$\max_{j\in[0,n-1]} \left|\widetilde{R}_{j} - \left(1 - \frac{j}{n}\right) R_{j,n}\right| = O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right),$$
$$\max_{j\in[0,n-1]} \left(1 - \frac{j}{n}\right)\left|\widetilde{\tau}_{j}^{2} - \tau_{j,n}^{2}\right| = O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{1/2}\right),$$

follow from Xiao and Wu (2011, Theorem 2). Note that these authors do not consider stationary sequences $\{u_{t,n}\}$ but their arguments carry over under Assumption R. Hence it suffices to study $\max_{j\in[0,\overline{p}_n]}\left|\widehat{R}_j-\widetilde{R}_j\right|$ and $\max_{j\in[0,\overline{p}_n]}\left|\widehat{\tau}_j^2-\widetilde{\tau}_j^2\right|$ since $\overline{p}_n/n=o\left(n^{-1/2}\right)$ under Assumption P. We then now show that $\max_{j\in[0,\overline{p}_n]}\left|\widehat{R}_j-\widetilde{R}_j\right|=O_{\mathbb{P}}\left(n^{-1/2}\right)$. Let $e_t=\widehat{u}_t-u_t$, so that

$$\widehat{R}_j=\frac{1}{n}\sum_{t=1}^{n-j}\left(u_t+e_t\right)\left(u_{t+j}+e_{t+j}\right)=\widetilde{R}_j+\frac{1}{n}\sum_{t=1}^{n-j}\left(u_te_{t+j}+e_tu_{t+j}\right)+\frac{1}{n}\sum_{t=1}^{n-j}e_te_{t+j}$$

with, by the Cauchy-Schwarz inequality, $\left|\sum_{t=1}^{n-j}e_te_{t+j}\right|/n\leq\sum_{t=1}^{n}e_t^2/n$ and, under Assumption M, for $\widehat{\mathfrak{r}}_t=\mathfrak{r}_t\left(\widehat{\theta}\right)$,

$$\frac{1}{n}\sum_{t=1}^{n-j}u_te_{t+j}=\left(\widehat{\theta}-\theta\right)\frac{1}{n}\sum_{t=1}^{n-j}u_tu_{t+j}^{(1)}+\frac{1}{2}\left(\widehat{\theta}-\theta\right)^2\frac{1}{n}\sum_{t=1}^{n-j}u_tu_{t+j}^{(2)}+\frac{1}{n}\sum_{t=1}^{n-j}u_t\widehat{\mathfrak{r}}_{t+j}.$$

Now, observe that Assumption M gives $\widehat{\theta}-\theta=O_{\mathbb{P}}\left(n^{-1/2}\right)$, $\max_{t\in[1,n]}\left|\widehat{\mathfrak{r}}_t\right|=o_{\mathbb{P}}\left(1/n\right)$ and

$$\frac{1}{n}\sum_{t=1}^{n}e_t^2\leq3\left(\widehat{\theta}-\theta\right)^2\frac{1}{n}\sum_{t=1}^{n}\left(u_t^{(1)}\right)^2+\frac{3}{4}\left(\widehat{\theta}-\theta\right)^4\frac{1}{n}\sum_{t=1}^{n}\left(u_t^{(1)}\right)^2+\frac{3}{n}\sum_{t=1}^{n}\left|\widehat{\mathfrak{r}}_t\right|=O_{\mathbb{P}}\left(\frac{1}{n}\right),$$

$$\max_{j\in[1,n]}\left|\frac{1}{n}\sum_{t=1}^{n-j}\left(u_t\widehat{\mathfrak{r}}_{t+j}+u_{t+j}\widehat{\mathfrak{r}}_t\right)\right|\leq\frac{2\max_{t\in[1,n]}\left|\widehat{\mathfrak{r}}_t\right|}{n}\sum_{t=1}^{n-j}\left|u_t\right|=o_{\mathbb{P}}\left(\frac{1}{n}\right).$$

This gives, uniformly in $j\in[1,n]$

$$\begin{aligned}\left|\widehat{R}_j-\widetilde{R}_j\right|&\leq\left|\widehat{\theta}-\theta\right|\left|\mathbb{E}\left[u_tu_{t+j}^{(1)}+u_{t+j}u_t^{(1)}\right]\right|\\&+\left|\widehat{\theta}-\theta\right|\left|\frac{1}{n}\sum_{t=1}^{n-j}\left(u_tu_{t+j}^{(1)}+u_{t+j}u_t^{(1)}-\mathbb{E}\left[u_tu_{t+j}^{(1)}+u_{t+j}u_t^{(1)}\right]\right)\right|+O_{\mathbb{P}}\left(\frac{1}{n}\right).\end{aligned}\tag{B.2}$$

It also follows from Assumption M and $\overline{p}_n=o\left(n^{1/2}\right)$ that $\left|\widehat{\theta}-\theta\right|\max_{j\in[1,n]}\left|\mathbb{E}\left[u_tu_{t+j}^{(1)}+u_{t+j}u_t^{(1)}\right]\right|=O_{\mathbb{P}}\left(1/n^{1/2}\right)$, $n\left(\widehat{\theta}-\theta\right)^2\sum_{j=0}^{\infty}\mathbb{E}^2\left[u_tu_{t+j}^{(1)}+u_{t+j}u_t^{(1)}\right]=O_{\mathbb{P}}\left(1\right)$, and for $A_t\left(j\right)=u_tu_{t+j}^{(1)}+$

$u_{t+j}u_t^{(1)} - \mathbb{E}\left[u_t u_{t+j}^{(1)} + u_{t+j}u_t^{(1)}\right]$

$$
\begin{aligned}
&\left|\widehat{\theta}-\theta\right| \max_{j\in[0,\overline{p}_n]} \left|\frac{1}{n}\sum_{t=1}^{n-j} A_t\left(j\right)\right| \leq O_{\mathbb{P}}\left(\frac{1}{n^{1/2}}\right)\sum_{j=0}^{\overline{p}_n}\left|\frac{1}{n}\sum_{t=1}^{n-j} A_t\left(j\right)\right| \\
&= O_{\mathbb{P}}\left(\frac{1}{n}\right) O_{\mathbb{P}}\left(\sum_{j=0}^{\overline{p}_n}\mathbb{E}^{1/2}\left[\left(\frac{1}{n^{1/2}}\sum_{t=1}^{n-j} A_t\left(j\right)\right)^2\right]\right) \\
&= O_{\mathbb{P}}\left(\frac{1}{n}\right) O_{\mathbb{P}}\left(\overline{p}_n \max_{j\in[0,\overline{p}_n]}\left[\left(\frac{1}{n^{1/2}}\sum_{t=1}^{n-j} A_t\left(j\right)\right)^2\right]\right) = O_{\mathbb{P}}\left(\frac{1}{n^{1/2}}\right),
\end{aligned}
$$

$$
\begin{aligned}
&n\sum_{j=0}^{n-1}\left(\widehat{\theta}-\theta\right)^2\left(\frac{1}{n}\sum_{t=1}^{n-j} A_t\left(j\right)\right)^2 \\
&= O_{\mathbb{P}}\left(1\right)\frac{1}{n} O_{\mathbb{P}}\left(\sum_{j=0}^{n-1}\mathbb{E}\left[\left(\frac{1}{n^{1/2}}\sum_{t=1}^{n-j} A_t\left(j\right)\right)^2\right]\right) \\
&= O_{\mathbb{P}}\left(1\right)\frac{1}{n} O_{\mathbb{P}}\left(n\max_{j\in[0,n]}\mathbb{E}\left[\left(\frac{1}{n^{1/2}}\sum_{t=1}^{n-j} A_t\left(j\right)\right)^2\right]\right) = O_{\mathbb{P}}\left(1\right).
\end{aligned}
$$

This gives $\max_{j\in[0,\overline{p}_n]}\left|\widehat{R}_j - \widetilde{R}_j\right| = O_{\mathbb{P}}\left(n^{-1/2}\right)$ and $\max_{p\in[0,n-1]} n\sum_{j=1}^{p}\left(\widehat{R}_j-\widetilde{R}_j\right)^2 = O_{\mathbb{P}}\left(1\right)$. The study of $\max_{j\in[0,\overline{p}_n]}\left|\widehat{\tau}_j^2 - \widetilde{\tau}_j^2\right|$ is similar. □

B.5. **Proof of Proposition A.2.** For the sake of brevity we assume that $\theta$ is unidimensional. Since $\widehat{R}_j^2 - \widetilde{R}_j^2 = \left(\widehat{R}_j - \widetilde{R}_j\right)^2 + 2\widetilde{R}_j\left(\widehat{R}_j - \widetilde{R}_j\right)$, Proposition A.2 is a direct consequence of Proposition A.1 and Lemma B.1 below.

**Lemma B.1.** *Assume that Assumptions K, M, P and R hold. Then*

$$
\max_{p\in[2,\overline{p}_n]} \frac{\left|n\sum_{j=1}^{n-1}\left(K^2(j/p)-K^2(j)\right)\widetilde{R}_j\left(\widehat{R}_j-\widetilde{R}_j\right)\right|}{\left(1+n\sum_{j=1}^{p} R_j^2\right)^{1/2}} = O_{\mathbb{P}}\left(1\right)
$$

*and* $n\sum_{j=1}^{n-1} K^2(j/p_n)\widetilde{R}_j\left(\widehat{R}_j-\widetilde{R}_j\right) = O_{\mathbb{P}}\left(\left(1+n\sum_{j=1}^{p_n} R_j^2\right)^{1/2}\right)$ *for any* $p_n = O(n^{1/2})$.

**Proof of Lemma B.1.** We just prove the first equality since the proof of the second is very similar. Define $\overline{R}_j = \mathbb{E}\left[\widetilde{R}_j\right] = (1 - j/n)R_j$. We have

$$\begin{aligned}
\left| n \sum_{j=1}^{n-1} K_{jp} \widetilde{R}_j \left( \widehat{R}_j - \widetilde{R}_j \right) \right| &\leq C_n(p) + D_n(p), \text{ where} \\
C_n(p) &= \left| n \sum_{j=1}^{n-1} K_{jp} R_j \left( \widehat{R}_j - \widetilde{R}_j \right) \right|, \\
D_n(p) &= \left| n \sum_{j=1}^{n-1} K_{jp} \left( \widetilde{R}_j - \overline{R}_j \right) \left( \widehat{R}_j - \widetilde{R}_j \right) \right|.
\end{aligned}$$

The Cauchy-Schwarz inequality and Assumption K gives

$$C_n(p) \leq C \left( n \sum_{j=1}^{p} R_j^2 \right)^{1/2} \left( n \sum_{j=1}^{p} \left( \widehat{R}_j - \widetilde{R}_j \right)^2 \right)^{1/2}.$$

Hence Proposition A.1 yields that $\max_{p\in[2,\overline{p}_n]} |C_n(p)/\left(n \sum_{j=1}^{p} R_j^2\right)^{1/2}| = O_{\mathbb{P}}(1)$. For $D_n(p)$, Assumptions K, M, (B.2) and $\widehat{\mathfrak{r}}_t = \mathfrak{r}_t\left(\widehat{\theta}\right)$ give

$$\begin{aligned}
\max_{p\in[2,\overline{p}_n]} D_n(p) \leq{}& O_{\mathbb{P}}(n^{-1/2}) \left( \max_{p\in[2,\overline{p}_n]} D_{1n}(p) + \max_{p\in[2,\overline{p}_n]} D_{2n}(p) \right) + O_{\mathbb{P}}(n^{-1}) \max_{p\in[2,\overline{p}_n]} D_{3n}(p) \\
&+ \left( \frac{1}{n} \sum_{t=1}^{n} e_t^2 + 2 \frac{\max_{t\in[1,n]} |\mathfrak{r}_t|}{n} \sum_{t=1}^{n} |u_t| \right) \max_{p\in[2,\overline{p}_n]} D_{4n}(p),
\end{aligned}$$

where $D_{1n}(p) = n \sum_{j=1}^{p} \left| \widetilde{R}_j - \overline{R}_j \right| \left| \mathbb{E}\left[ u_t u_{t+j}^{(1)} + u_{t+j} u_t^{(1)} \right] \right|$,

$$\begin{aligned}
D_{2n}(p) &= n \sum_{j=1}^{p} \left| \widetilde{R}_j - \overline{R}_j \right| \left| \frac{1}{n} \sum_{t=1}^{n-j} \left( u_t u_{t+j}^{(1)} + u_{t+j} u_t^{(1)} - \mathbb{E}\left[ u_t u_{t+j}^{(1)} + u_{t+j} u_t^{(1)} \right] \right) \right|, \\
D_{3n}(p) &= n \sum_{j=1}^{p} \left| \widetilde{R}_j - \overline{R}_j \right| \left| \frac{1}{n} \sum_{t=1}^{n-j} \left( u_t u_{t+j}^{(2)} + u_{t+j} u_t^{(2)} \right) \right|, \\
D_{4n}(p) &= n \sum_{j=1}^{p} \left| \widetilde{R}_j - \overline{R}_j \right|.
\end{aligned}$$

By Assumption K and M and by Lemma A.4, we have

$$\mathbb{E}\left[\max_{p\in[2,\overline{p}_n]} D_{1n}(p)\right] \leq Cn\sum_{j=1}^{\overline{p}_n} \mathrm{Var}^{1/2}\left(\widetilde{R}_j\right)\left|\mathbb{E}\left[u_t u_{t+j}^{(1)} + u_{t+j}u_t^{(1)}\right]\right| \leq Cn^{1/2},$$

$$\begin{aligned}\mathbb{E}\left[\max_{p\in[2,\overline{p}_n]} D_{2n}(p)\right] &\leq Cn^{1/2}\sum_{j=1}^{\overline{p}_n} \mathrm{Var}^{1/2}\left(\widetilde{R}_j\right)\\ &\quad\times \mathbb{E}^{1/2}\left[\left|\frac{1}{n^{1/2}}\sum_{t=1}^{n}\left(u_t u_{t+j}^{(1)} + u_{t+j}u_t^{(1)} - \mathbb{E}\left[u_t u_{t+j}^{(1)} + u_{t+j}u_t^{(1)}\right]\right)\right|^2\right]\\ &\leq C\overline{p}_n,\end{aligned}$$

$$\mathbb{E}\left[\max_{p\in[2,\overline{p}_n]} D_{3n}(p)\right] \leq Cn\sum_{j=1}^{\overline{p}_n} \mathrm{Var}^{1/2}\left(\widetilde{R}_j\right)\mathbb{E}^{1/2}\left[\left|\frac{1}{n}\sum_{t=1}^{n}\left(u_t u_{t+j}^{(2)} + u_{t+j}u_t^{(2)}\right)\right|^2\right] \leq C\overline{p}_n n^{1/2},$$

$$\mathbb{E}\left[\max_{p\in[2,\overline{p}_n]} D_{4n}(p)\right] \leq Cn\sum_{j=1}^{\overline{p}_n} \mathbb{E}\left[\left|\widetilde{R}_j - \overline{R}_j\right|\right] \leq Cn\sum_{j=1}^{\overline{p}_n} \mathrm{Var}^{1/2}\left(\widetilde{R}_j\right) \leq Cn^{1/2}\overline{p}_n.$$

The Markov inequality gives us the stochastic orders of magnitude of the four maxima in the bound for $\max_{p\in[2,\overline{p}_n]} D_n(p)$. Since $\overline{p}_n = O\left(n^{1/2}\right)$ by Assumption P, $\max_{t\in[1,n]}|\widehat{\mathfrak{r}}_t| = o_{\mathbb{P}}\left(1/n\right)$ and $n^{-1}\sum_{t=1}^{n} e_t^2 = O_{\mathbb{P}}(n^{-1})$ by Assumption M, we have $\max_{p\in[2,\overline{p}_n]}|D_n(p)| = O_{\mathbb{P}}\left(1 + \frac{\overline{p}_n}{n^{1/2}}\right) = O_{\mathbb{P}}\left(1\right)$. This together with $\max_{p\in[2,\overline{p}_n]}|C_n(p)/\left(n\sum_{j=1}^{p} R_j^2\right)^{1/2}| = O_{\mathbb{P}}\left(1\right)$ shows that the Lemma is proved. □

B.6. **Proof of Proposition A.3.** The proof of Proposition A.3 is long and divided in three steps. In the two first steps, we focus on observed variables. In the first step, we approximate the sample covariance $\widetilde{R}_j$ by a martingale counterpart $\sum_{t=1}^{n} D_{jt}/n$, $j\in[1,\overline{p}_n]$, as in Shao (2011b), see the notations below and Lemmas B.2, B.3. and B.4. The second step deals with the deviation probability of

$$\frac{n\sum_{j=1}^{p}\left(\frac{1}{n}\sum_{t=j+1}^{n} D_{jt}\right)^2\left(K^2\left(j/p\right) - K^2\left(1\right)\right) - \sigma^4 E_{\Delta}\left(p\right)}{\sigma^4 V_{\Delta}\left(p\right)}$$

which is approximated with some Gaussian counterparts through the Lindeberg technique, see Lemma B.5. The third step concludes and explicitly deals with the case of residuals thanks to Propositions A.1 and A.2.

Let us now introduce additional notations. Let $\mathcal{F}_k$ be the sigma field generated by $e_k, e_{k-1}, \ldots$. Define $\mathbf{P}_t[Z] = \mathbb{E}[Z|\mathcal{F}_t] - \mathbb{E}[Z|\mathcal{F}_{t-1}]$. Wu (2007, Proposition 3) establishes that $\|\mathbf{P}_t[u_{t+k}]\|_a \leq \delta_a(k)$ and Shao (2011b) has shown that

$$\|\mathbf{P}_0[u_k u_{k-j}]\|_a \leq 2\|u_k\|_{2a}(\delta_{2a}(k) + \delta_{2a}(k-j)\mathbb{I}(j \leq k)), \tag{B.3}$$

which is smaller than $4\|u_k\|_{2a}\delta_{2a}(k-j)$ when $j \leq k$. Define now the vector of martingale difference $D_t = \left[D_{1t}, \ldots, D_{\overline{p}_n t}\right]'$ with

$$D_{jt} = \sum_{k=t}^{\infty} \mathbf{P}_t[u_k u_{k-j}]$$

which converges a.s. and satisfies $\mathbb{E}[D_{jt}|\mathcal{F}_{t-1}] = 0$, $\max_j \mathbb{E}[|D_{jt}|^a] < \infty$, provided $\|u_t\|_{2a} < \infty$ and $\sum_{k=0}^{\infty}\delta_{2a}(k) < \infty$. Consider the martingale $M_j = M_{jn} = \sum_{t=j+1}^{n} D_{jt}$ which is an approximation of $\widetilde{R}_j$. Shao (Lemma A.1, 2011b) gives under Assumption R and for any $\mathfrak{a} \in [1, 6a]$,

$$\left(\mathbb{E}^{\frac{1}{\mathfrak{a}}}\left[\left|\sum_{t=j+1}^{n} u_t u_{t-j} - M_j\right|^{\mathfrak{a}}\right]\right)^2 \leq C. \tag{B.4}$$

We shall also use a $\mathfrak{p}$-dependent version of $D_t$, denoted $D_t^{t-\mathfrak{p}+1}$, with entries

$$D_{jt}^{t-\mathfrak{p}+1} = \mathbb{E}[D_{jt}|e_t, \ldots, e_{t-\mathfrak{p}+1}] = \sum_{k=t}^{\infty}\mathbf{P}_t'[u_k u_{k-j}], \text{ where} \tag{B.5}$$

$$\mathbf{P}_t'[Z] = \mathbf{P}_t^{t-\mathfrak{p}+1}[Z] = \mathbb{E}[Z|e_t, \ldots, e_{t-\mathfrak{p}+1}] - \mathbb{E}[Z|e_{t-1}, \ldots, e_{t-\mathfrak{p}+1}].$$

Arguing as in Shao (2011b, Lemma A.2-(iii)) gives

$$\left\|D_{jt} - D_{jt}^{t-\mathfrak{p}+1}\right\|_{\mathfrak{a}} \leq C\|u_t\|_{2\mathfrak{a}}\Theta_{2\mathfrak{a}}(\mathfrak{p}-j), \quad \text{for all } j \in [1, \mathfrak{p}]. \tag{B.6}$$

B.6.1. *Martingale approximation and preliminary lemmas.* An important property of $D_t$ and $D_t^{t-\mathfrak{p}+1}$ is as follows.

**Lemma B.2.** *Suppose Assumption K and R hold. Let $K_{jp}$ be as in (B.1). Then for any $p \leq \mathfrak{p}$, $t$, and any $s \leq t-\mathfrak{p}$, $\left\|\sum_{j=1}^{p} K_{jp} D_{js} D_{jt}^{t-\mathfrak{p}+1}\right\|_{3a} \leq C p^{1/2}$.*

**Proof of Lemma B.2**. We have

$$\begin{aligned}
&\left\|\sum_{j=1}^{p} K_{jp} D_{js} D_{jt}^{t-\mathfrak{p}+1}\right\|_{3a} \\
&\quad= \left\|\sum_{j=1}^{p} K_{jp} \sum_{k_1=0}^{\infty} \mathbf{P}_s\left[u_{s+k_1} u_{s+k_1-j}\right] \sum_{k_2=0}^{\infty} \mathbf{P}_t'\left[u_{t+k_2} u_{t+k_2-j}\right]\right\|_{3a} \\
&\quad\leq \left\|\sum_{j=1}^{p} K_{jp} \sum_{k_1=0}^{j-1} \mathbf{P}_s\left[u_{s+k_1} u_{s+k_1-j}\right] \sum_{k_2=0}^{j-1} \mathbf{P}_t'\left[u_{t+k_2} u_{t+k_2-j}\right]\right\|_{3a} \qquad \text{(B.7)} \\
&\quad+ \left\|\sum_{j=1}^{p} K_{jp} \sum_{k_1=0}^{j-1} \mathbf{P}_s\left[u_{s+k_1} u_{s+k_1-j}\right] \sum_{k_2=j}^{\infty} \mathbf{P}_t'\left[u_{t+k_2} u_{t+k_2-j}\right]\right\|_{3a} \qquad \text{(B.8)} \\
&\quad+ \left\|\sum_{j=1}^{p} K_{jp} \sum_{k_1=j}^{\infty} \mathbf{P}_s\left[u_{s+k_1} u_{s+k_1-j}\right] \sum_{k_2=0}^{j-1} \mathbf{P}_t'\left[u_{t+k_2} u_{t+k_2-j}\right]\right\|_{3a} \qquad \text{(B.9)} \\
&\quad+ \left\|\sum_{j=1}^{p} K_{jp} \sum_{k_1=j}^{\infty} \mathbf{P}_s\left[u_{s+k_1} u_{s+k_1-j}\right] \sum_{k_2=j}^{\infty} \mathbf{P}_t'\left[u_{t+k_2} u_{t+k_2-j}\right]\right\|_{3a}. \qquad \text{(B.10)}
\end{aligned}$$

We have for (B.7)

$$\begin{aligned}
\text{(B.7)} &= \left\|\sum_{j=1}^{p} K_{jp} \sum_{k_1=0}^{p-1} \mathbb{I}\left(k_1<j\right) u_{s+k_1-j} \mathbf{P}_s\left[u_{s+k_1}\right] \sum_{k_2=0}^{p-1} \mathbb{I}\left(k_2<j\right) u_{t+k_2-j} \mathbf{P}_t'\left[u_{t+k_2}\right]\right\|_{3a} \\
&= \left\|\sum_{k_1=0}^{p-1} \sum_{k_2=0}^{p-1}\left(\sum_{j=k_1 \vee k_2}^{p-1} K_{jp} u_{s+k_1-j} u_{t+k_2-j}\right) \mathbf{P}_s\left[u_{s+k_1}\right] \mathbf{P}_t'\left[u_{t+k_2}\right]\right\|_{3a} \\
&\leq \sum_{k_1=0}^{p-1} \sum_{k_2=0}^{p-1}\left\|\sum_{j=k_1 \vee k_2}^{p-1} K_{jp} u_{s+k_1-j} u_{t+k_2-j}\right\|_{6a} \delta_{12a}\left(k_1\right) \delta_{12a}\left(k_2\right),
\end{aligned}$$

using $\left\|\mathbf{P}_t'\left[u_{t+k_2}\right]\right\|_{12a} \leq \left\|\mathbf{P}_t\left[u_{t+k_2}\right]\right\|_{12a} = \delta_{12a}\left(k_2\right)$. Now (B.4) and the Burkholder inequality give

$$\left\|\sum_{j=k_1\vee k_2}^{p-1} K_{jp} u_{s+k_1-j} u_{t+k_2-j}\right\|_{6a} \leq \left\|\sum_{j=k_1\vee k_2}^{p-1} K_{jp} D_{t+k_2-j,t-s+k_2-k_1}\right\|_{6a}$$
$$+\left\|\sum_{j=k_1\vee k_2}^{p-1} K_{jp}\left(u_{s+k_1-j} u_{t+k_2-j} - D_{t+k_2-j,t-s+k_2-k_1}\right)\right\|_{6a} \leq C p^{1/2}.$$

Hence (B.7) is smaller than $Cp^{1/2}$. For (B.8), we have since $\left\{u_{s+k_1-j}, j \in [1, k_1]\right\}$ and $\left\{\mathbf{P}_t'\left[u_{t+k_2} u_{t+k_2-j}\right], j \in [1, k_1], k_2 \geq 0\right\}$ are independent,

$$\text{(B.8)} = \left\|\sum_{k_1=0}^{p-1}\sum_{k_2=0}^{\infty}\left(\sum_{j=k_1}^{p-1} K_{jp} u_{s+k_1-j}\mathbf{P}_t'\left[u_{t+k_2+j} u_{t+k_2}\right]\right)\mathbf{P}_s\left[u_{s+k_1}\right]\right\|_{3a}$$
$$\leq \sum_{k_1=0}^{p-1}\sum_{k_2=0}^{\infty}\left\|\sum_{j=k_1}^{p-1} K_{jp} u_{s+k_1-j}\mathbf{P}_t'\left[u_{t+k_2+j} u_{t+k_2}\right]\right\|_{6a}\delta_{6a}\left(k_1\right).$$

Let $d_t = \sum_{k=t}^{\infty}\mathbf{P}_t\left[u_k\right]$ be the martingale difference approximation of $u_t$, see Wu (2007). Now, since $\left\{u_{s+k_1-j}, d_{s+k_1-j,} j \in [1, k_1]\right\}$ and $\left\{\mathbf{P}_t'\left[u_{t+k_2} u_{t+k_2-j}\right], j \in [1, k_1], k_2 \geq 0\right\}$ are independent, arguing as in the proof of Theorem 1 in Wu (2007), (B.4) and the Burkholder inequality give

$$\left\|\sum_{j=k_1}^{p-1} K_{jp} u_{s+k_1-j}\mathbf{P}_t'\left[u_{t+k_2+j} u_{t+k_2}\right]\right\|_{6a}^2$$
$$\leq 2\left\|\sum_{j=k_1}^{p-1} K_{jp} d_{s+k_1-j}\mathbf{P}_t'\left[u_{t+k_2+j} u_{t+k_2}\right]\right\|_{6a}^2 + 2\left\|\sum_{j=k_1}^{p-1} K_{jp}\left(u_{s+k_1-j} - d_t\right)\mathbf{P}_t'\left[u_{t+k_2+j} u_{t+k_2}\right]\right\|_{6a}^2$$
$$\leq C\left\|\sum_{j=k_1}^{p-1} K_{jp} d_{s+k_1-j}^2\left(\mathbf{P}_t'\left[u_{t+k_2+j} u_{t+k_2}\right]\right)^2\right\|_{3a} + C\left\|\mathbf{P}_t'\left[u_{t+k_2+j} u_{t+k_2}\right]\right\|_{6a}^2 \leq C k_1 \delta_{6a}^2\left(k_2\right).$$

Hence Assumption R gives (B.8)$\leq \sum_{k_1=0}^{p-1}\sum_{k_2=0}^{\infty} k_1 \delta_{6a}^2\left(k_2\right)\delta_{6a}\left(k_1\right) \leq C$.

For (B.9), observe first that (B.4) gives

$$\begin{aligned}
\text{(B.9)} &= \left\| \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{p-1} \sum_{j=1}^{p} K_{jp} \mathbb{I}\left(j \le k_1\right) \mathbf{P}_s \left[u_{s+k_1} u_{s+k_1-j}\right] \mathbb{I}\left(k_2 < j\right) \mathbf{P}_t' \left[u_{t+k_2} u_{t+k_2-j}\right] \right\|_{3a} \\
&\le \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{p-1} \sum_{j=k_2}^{p} \mathbb{I}\left(j \le k_1\right) \delta_{6a}\left(k_1 - j\right) \left\| \mathbf{P}_t' \left[u_{t+k_2} u_{t+k_2-j}\right] \right\|_{6a} \\
&\le \left( \sum_{k_1=0}^{\infty} \delta_{6a}\left(k_1\right) \right) \times \sum_{k_2=0}^{p-1} \sum_{j=k_2}^{p} \left\| \mathbf{P}_t' \left[u_{t+k_2} u_{t+k_2-j}\right] \right\|_{6a}.
\end{aligned}$$

Since $\overline{u}_{t+k_2-j}^{t}$ is independent of $e_t, \ldots, e_{t-\mathfrak{p}+1}$ and $\mathbf{P}_t\left[u_{t+k_2}\right]$,

$$\begin{aligned}
\left\| \mathbf{P}_t' \left[u_{t+k_2} u_{t+k_2-j}\right] \right\|_{6a} &\le \left\| \underbrace{\mathbb{E}\left[ \overline{u}_{t+k_2-j}^{t} \mathbf{P}_t\left[u_{t+k_2}\right] \middle| e_t, \ldots, e_{t-\mathfrak{p}+1} \right]}_{0} \right\|_{6a} \\
&+ \left\| \mathbb{E}\left[ \left(u_{t+k_2-j} - \overline{u}_{t+k_2-j}^{t}\right) \mathbf{P}_t\left[u_{t+k_2}\right] \middle| e_t, \ldots, e_{t-\mathfrak{p}+1} \right] \right\|_{6a} \\
&\le \left\| u_{t+k_2-j} - \overline{u}_{t+k_2-j}^{t} \right\|_{12a} \left\| \mathbf{P}_t\left[u_{t+k_2}\right] \right\|_{12a} \le \Theta_{12a}\left(k_2 - j\right) \delta_{12a}\left(k_2\right).
\end{aligned} \tag{B.11}$$

Substituting gives that (B.9)$\le C \sum_{k_2=0}^{p-1} \sum_{j=k_2}^{p} \Theta_{12a}\left(k_2 - j\right) \delta_{12a}\left(k_2\right) \le C$.

For (B.10), (B.3) and (B.11) give

$$\begin{aligned}
\text{(B.10)} &\le C \sum_{j=1}^{p} \left( \sum_{k_1=j}^{\infty} \left\| \mathbf{P}_s \left[u_{s+k_1} u_{s+k_1-j}\right] \right\|_{6a} \right) \sum_{k_2=j}^{\infty} \left\| \mathbf{P}_t' \left[u_{t+k_2} u_{t+k_2-j}\right] \right\|_{6a} \\
&\le C \sum_{j=1}^{p} \left( \sum_{k_1=j}^{\infty} \delta_{6a}\left(k_1 - j\right) \right) \sum_{k_2=j}^{\infty} \Theta_{12a}\left(k_2 - j\right) \delta_{12a}\left(k_2\right) \le C.
\end{aligned}$$

Hence substituting gives $\left\| \sum_{j=1}^{p} K_{jp} D_{js} D_{jt}^{t-\mathfrak{p}+1} \right\|_{3a} \le C p^{1/2}$. □

We now define a suitable sequence of Gaussian vector. Let $2\overline{p}_n \le \ell \le 3\overline{p}_n$ be an integer number. Consider a sequence of independent centered Gaussian vectors $\eta_t = \left[\eta_{1t}, \ldots, \eta_{\overline{p}_n t}\right]'$ with

$$\mathbb{E}\left[\eta_{j_1 t} \eta_{j_2 t}\right] = \mathbb{E}\left[ D_{j_1 t}^{t-\ell+1} D_{j_2 t}^{t-\ell+1} \right]. \tag{B.12}$$

We shall also assume that $\{\eta_t\}$ and $\{e_t\}$ are independent.

**Lemma B.3.** *Let $\{\eta_t\}$ be as in (B.12) and suppose Assumption R holds. Then for all $p \in [1, \overline{p}_n]$ and $t, s \in [1, n]$,*

$$\sum_{j_1 \neq j_2 \in [1, \overline{p}_n]} |\mathrm{Cov}\,(\eta_{j_1 t}, \eta_{j_2 t})| \leq C \ \ and \ \sum_{j=1}^{\overline{p}_n} \left|\mathrm{Var}\,(\eta_{jt}) - \sigma^4\right| \leq C,$$
$$\left|\sum_{j=1}^{p} \left(1 - \frac{j}{n}\right) K_{jp} \left(\mathrm{Var}\,(\eta_{jt}) - \sigma^4\right)\right| \leq C,$$
$$\left|\left(2 \sum_{j=1}^{p} \left(1 - \frac{j}{n}\right)^2 K_{jp}^2 \,\mathrm{Var}^2\,(\eta_{jt})\right)^{1/2} - \sigma^4 V_\Delta\,(p)\right| \leq C,$$
$$\mathrm{Var}\left(\frac{1}{p^{1/2}} \sum_{j=1}^{p} K_{jp} D_{js} \eta_{jt} \,| D_s\right) \leq \frac{C}{p} \sum_{j=1}^{p} K_{jp}^2 D_{js}^2.$$

**Proof of Lemma B.3.** (B.4) gives for all $j_1$, $j_2$,

$$\mathrm{Cov}\,(D_{j_1 t}, D_{j_2 t}) = \lim_{n \to \infty} \mathrm{Cov}\left(\frac{\sum_{t=j_1+1}^{n} u_t u_{t-j_1}}{(n-j_1)^{1/2}}, \frac{\sum_{t=j_2+1}^{n} u_t u_{t-j_2}}{(n-j_2)^{1/2}}\right) = \sum_{k=-\infty}^{\infty} \mathbb{E}\left[u_0 u_{j_1} u_k u_{k+j_2}\right],$$

see also Lemma A.2 in Shao (2011b), provided $\sum_{k=-\infty}^{\infty} |\mathbb{E}\left[u_0 u_{j_1} u_k u_{k+j_2}\right]| < \infty$ as shown below. (B.6) and (B.12) give

$$\max_{j_1, j_2 \in [0, \overline{p}_n]} \left|\mathrm{Cov}\,(\eta_{j_1 t}, \eta_{j_2 t}) - \sum_{k=-\infty}^{\infty} \mathbb{E}\left[u_0 u_{j_1} u_k u_{k+j_2}\right]\right| \leq C \Theta_{12a}\,(\overline{p}_n)\,. \tag{B.13}$$

Now relation between cumulants and moments in Brillinger (2001) and Theorem B.1 gives absolute summability of the 4th moments. Hence $\Theta_{12a}\,(\overline{p}_n) = O(\overline{p}_n^{-6})$ gives the first bound of the Lemma. For the second and the third bound, observe that under the null

$$\left|\sum_{k=-\infty}^{\infty} \mathbb{E}\left[u_0 u_j u_k u_{k+j}\right] - \sigma^4\right| \leq \left|\mathbb{E}\left[u_0^2 u_j^2\right] - \mathbb{E}\left[u_0^2\right] \mathbb{E}\left[u_j^2\right]\right| + 2\left|\sum_{k=1}^{\infty} \mathbb{E}\left[u_0 u_j u_k u_{k+j}\right]\right|.$$

$\left|\mathbb{E}\left[u_0^2 u_j^2\right] - \mathbb{E}\left[u_0^2\right]\mathbb{E}\left[u_j^2\right]\right| \leq C\Theta_{12a}\left(j\right) = O\left(j^{-6}\right)$ and absolute summability of the 4th moments gives the second bound. This also gives the fourth one since

$$\begin{aligned}
&\left|\left(2\sum_{j=1}^{p}\left(1-\frac{j}{n}\right)^2 K_{jp}^2 \operatorname{Var}^2\left(\eta_{jt}\right)\right)^{1/2} - \sigma^4 V_\Delta\left(p\right)\right| \\
&\leq \left(2\sum_{j=1}^{p}\left(1-\frac{j}{n}\right)^2 K_{jp}^2 \left(\operatorname{Var}\left(\eta_{jt}\right)-\sigma^4\right)^2\right)^{1/2} \\
&\leq 2^{1/2}\left|\sum_{j=1}^{p}\left(1-\frac{j}{n}\right) K_{jp}\left(\operatorname{Var}\left(\eta_{jt}\right)-\sigma^4\right)\right| \leq C.
\end{aligned}$$

For the last one, observe first that

$$\sum_{1\leq j_1<j_2\leq \overline{p}_n}\left|\operatorname{Cov}\left(\eta_{j_1 t},\eta_{j_2 t}\right)\right|^2 \leq \left(\sum_{1\leq j_1<j_2\leq \overline{p}_n}\left|\operatorname{Cov}\left(\eta_{j_1 t},\eta_{j_2 t}\right)\right|\right)^2 < \infty$$

by Theorem B.1 since the 2th cumulants are the covariance. This gives, for any $z = \left[z_1,\ldots,z_{\overline{p}_n}\right]'$,

$$\begin{aligned}
\operatorname{Var}\left(z'\eta\right) &= z'\mathbb{E}\left[\eta\eta'\right]z \leq \sum_{j=1}^{\overline{p}_n}\operatorname{Var}\left(\eta_{jt}\right) z_j^2 + 2\sum_{1\leq j_1<j_2\leq \overline{p}_n}\left|\operatorname{Cov}\left(\eta_{j_1 t},\eta_{j_2 t}\right)\right|\left|z_{j_1}\right|\left|z_{j_2}\right| \\
&\leq Czz' + 2\left(\sum_{1\leq j_1<j_2\leq \overline{p}_n}\left|\operatorname{Cov}\left(\eta_{j_1 t},\eta_{j_2 t}\right)\right|^2\right)^{1/2}\left(\sum_{1\leq j_1<j_2\leq \overline{p}_n} z_{j_1}^2 z_{j_2}^2\right)^{1/2} \\
&\leq Cz'z.
\end{aligned}$$

Hence $\operatorname{Var}\left(\sum_{j=1}^{p} K_{jp} D_{js}\eta_{jt}\left|D_s\right.\right) \leq C\left(\sum_{j=1}^{p} K_{jp}^2 D_{js}^2\right)^{1/2}$ since $\{D_t\}$ and $\{\eta_t\}$ are independent. □

B.6.2. *The deviation probability of the maximum of Proposition A.3.* The proof is based on a smooth approximation of the maximum of real numbers $x_1,\ldots,x_{\overline{p}_n}$. Consider an increasing and three times continuously differentiable real function $f$ with

$$\lim_{x\to-\infty} f\left(x\right) = 1, \quad f\left(x\right) = x \text{ for } x\geq 2, \quad \max_{i=1,2,3}\sup_x\left|f^{(i)}\left(x\right)\right| < \infty. \tag{B.14}$$

Let $e = e_n \to \infty$ with $\ln(\overline{p}_n)/e = o(1)$. Then $\max_{p\in[1,\overline{p}_n]}\{f(x_p)\} \leq \left(\sum_{p=1}^{\overline{p}_n} f^e(x_p)\right)^{1/e} \leq \overline{p}_n^{1/e}\max_{p\in[1,\overline{p}_n]}\{f(x_p)\}$ gives that

$$\left(\sum_{p=1}^{\overline{p}_n} f^e(x_p)\right)^{1/e} = \left(1 + O\left(\frac{\ln \overline{p}_n}{e}\right)\right)\max_{p\in[1,\overline{p}_n]}\{f(x_p)\}. \tag{B.15}$$

We will first find a suitable approximation for the distribution of

$$\mathcal{M} = \left(\sum_{p=1}^{\overline{p}_n} f^e(\check{s}_p)\right)^{1/e} \text{ where } \check{S}_p = n\sum_{j=1}^{p} K_{jp}\left(\frac{M_{jn}}{n}\right)^2, \quad \check{s}_p = \frac{\check{S}_p - \sigma^4 E_\Delta(p)}{\sigma^4 V_\Delta(p)}. \tag{B.16}$$

Define, for $\eta = \left[\eta_1, \ldots, \eta_{\overline{p}_n}\right]'$ and $x \in [0,1]$,

$$\begin{aligned}
M_{jt}(x;\eta) &= \sum_{s=j+1}^{t-1} D_{js} + x\eta_j + \sum_{s=t+1}^{n}\eta_{js}, \quad R_{jt}(x;\eta) = \frac{M_{jt}(x;\eta)}{n} \\
\check{s}_{pt}(x;\eta) &= \frac{n\sum_{j=1}^{p} K_{jp}R_{jt}^2(x;\eta) - \sigma^4 E_\Delta(p)}{\sigma^4 V_\Delta(p)}, \quad \Sigma_t(x;\eta) = f(\check{s}_{pt}(x;\eta)), \\
\mathcal{M}_t(x;\eta) &= \left(\sum_{p=1}^{\overline{p}_n}\Sigma_t^e(x;\eta)\right)^{\frac{1}{e}}, \quad \mathcal{M}_t(\eta) = \mathcal{M}_t(1;\eta),
\end{aligned} \tag{B.17}$$

and

$$\begin{aligned}
\check{s}_{pt}^{(1)}(x;\eta) &= \frac{d\check{s}_{pt}(x;\eta)}{dx} = \frac{2\sum_{j=1}^{p} K_{jp}\left(\sum_{s=j+1}^{t-1} D_{js} + x\eta_j + \sum_{s=t+1}^{n}\eta_{js}\right)\eta_j}{n\sigma^4 V_\Delta(p)}, \\
\check{s}_{pt}^{(2)}(x;\eta) &= \frac{d_{pt}^2\check{s}(x;\eta)}{dx^2} = \frac{2\sum_{j=1}^{p} K_{jp}\eta_j^2}{n\sigma^4 V_\Delta(p)}, \\
\Sigma_{pt}^{(1)}(x;\eta) &= f^{(1)}(\check{s}_{pt}(x;\eta))\,\check{s}_{pt}^{(1)}(x;\eta), \\
\Sigma_{pt}^{(2)}(x;\eta) &= f^{(2)}(\check{s}_{pt}(x;\eta))\left(\check{s}_{pt}^{(1)}(x;\eta)\right)^2 + f^{(1)}(\check{s}_{pt}(x;\eta))\,\check{s}_{pt}^{(2)}(x;\eta), \\
\Sigma_{pt}^{(3)}(x;\eta) &= f^{(3)}(\check{s}_{pt}(x;\eta))\left(\check{s}_{pt}^{(1)}(x;\eta)\right)^3 + 3f^{(2)}(\check{s}_{pt}(x;\eta))\,\check{s}_{pt}^{(1)}(x;\eta)\,\check{s}_{pt}^{(2)}(x;\eta).
\end{aligned}$$

We first bound the moments of $\Sigma_{pt}^{(1)}(x;\eta)$, $\Sigma_{pt}^{(2)}(x;\eta)$ and $\Sigma_{pt}^{(3)}(x;\eta)$ when $\eta$ is set to $D_t$ or $\eta_t$.

**Lemma B.4.** *Under Assumption R and if* $\overline{p}_n = O\left(n^{1/2}\right)$*, we have uniformly in* $p \in [1, \overline{p}_n]$*,* $x \in [0,1]$ *and* $t = 1, \ldots, n$*,*

$$\max\left\{\left\|\Sigma_{pt}^{(1)}(x; D_t)\right\|_{3a}, \left\|\Sigma_{pt}^{(1)}(x; \eta_t)\right\|_{3a}\right\} \leq \frac{C}{n^{1/2}}, \tag{B.18}$$

$$\max\left\{\left\|\Sigma_{pt}^{(2)}(x; D_t)\right\|_{3a/2}, \left\|\Sigma_{pt}^{(2)}(x; \eta_t)\right\|_{3a/2}\right\} \leq \frac{Cp^{1/2}}{n}, \tag{B.19}$$

$$\max\left\{\left\|\Sigma_{pt}^{(3)}(x; D_t)\right\|_{a}, \left\|\Sigma_{pt}^{(3)}(x; \eta_t)\right\|_{a}\right\} \leq \frac{Cp^{1/2}}{n^{3/2}}. \tag{B.20}$$

**Proof of Lemma B.4.** (B.14) gives

$$\begin{aligned}
&\left|\Sigma_{pt}^{(1)}(x; \eta)\right| \leq C\left|\check{s}_{pt}^{(1)}(x; \eta)\right|, \quad \left|\Sigma_{pt}^{(2)}(x; \eta)\right| \leq C\left(\left(\check{s}_{pt}^{(1)}(x; \eta)\right)^2 + \left|\check{s}_{pt}^{(2)}(x; \eta)\right|\right), \\
&\left|\Sigma_{pt}^{(3)}(x; \eta)\right| \leq C\left|\check{s}_{pt}^{(1)}(x; \eta)\right|\left(\left(\check{s}_{pt}^{(1)}(x; \eta)\right)^2 + \left|\check{s}_{pt}^{(2)}(x; \eta)\right|\right).
\end{aligned} \tag{B.21}$$

(B.21) shows that the lemma directly follows from

$$\max\left\{\left\|\check{s}_{pt}^{(1)}(x; D_t)\right\|_{3a}, \left\|\check{s}_{pt}^{(1)}(x; \eta_t)\right\|_{3a}\right\} \leq \frac{C}{n^{1/2}}, \tag{B.22}$$

$$\max\left\{\left\|\check{s}_{pt}^{(2)}(x; D_t)\right\|_{3a/2}, \left\|\check{s}_{pt}^{(2)}(x; \eta_t)\right\|_{3a/2}\right\} \leq \frac{Cp^{1/2}}{n}. \tag{B.23}$$

(B.23) directly follow from the triangular inequality. For (B.22), we first bound $\left\|\check{s}_{pt}^{(1)}(x; D_t)\right\|_{3a}$. We have

$$\left\|\check{s}_{pt}^{(1)}(x; D_t)\right\|_{3a} \leq C\left\|\frac{\sum_{s=1}^{t-1}\left(\sum_{j=1}^{p} K_{jp} D_{js} D_{jt}\right)}{np^{1/2}}\right\|_{3a} \tag{B.24}$$

$$+ C\left\|\frac{\sum_{j=1}^{p} K_{jp} D_{jt}^2}{np^{1/2}}\right\|_{3a} + C\left\|\frac{\sum_{s=t+1}^{n}\left(\sum_{j=1}^{p} K_{jp} D_{jt} \eta_{js}\right)}{np^{1/2}}\right\|_{3a}. \tag{B.25}$$

We have, for the first item (B.24)

$$
\begin{aligned}
\text{(B.24)} \leq & \left\| \frac{\sum_{j=1}^{p} D_{jt} \sum_{s=1}^{t-\mathfrak{p}} K_{jp} D_{js}}{np^{1/2}} \right\|_{3a} + \left\| \frac{\sum_{s=t-\mathfrak{p}+1}^{t-1} D_{jt} \sum_{j=1}^{p} K_{jp} D_{js}}{np^{1/2}} \right\|_{3a} \\
\leq & \left\| \frac{\sum_{j=1}^{p} D_{jt} \sum_{s=1}^{t-\mathfrak{p}} K_{jp} D_{js}}{np^{1/2}} \right\|_{3a} + \frac{1}{np^{1/2}} \sum_{j=1}^{p} \left\| K_{jp} D_{jt} \right\|_{6a} \left\| \sum_{s=t-\mathfrak{p}+1}^{t-1} D_{js} \right\|_{6a} \\
\leq & \left\| \frac{\sum_{s=1}^{t-\mathfrak{p}} K_{jp} \sum_{j=1}^{p} D_{jt} D_{js}}{np^{1/2}} \right\|_{3a} + \frac{Cp^{1/2}\mathfrak{p}^{1/2}}{n},
\end{aligned}
$$

where $\mathfrak{p} \geq p$ and by the Burkholder inequality. Now let $\widetilde{D}_{jt} = D_{jt}^{t-\mathfrak{p}+1}$ be as in (B.5). Since $\sum_{j=1}^{p} K_{jp} D_{js} \widetilde{D}_{jt}$ is a martingale difference given $e_t, \ldots, e_{t-\mathfrak{p}+1}$, (B.6), the Burkholder and triangular inequalities, Lemma B.2 give

$$
\begin{aligned}
& \left\| \frac{\sum_{j=1}^{p} \sum_{s=1}^{t-\mathfrak{p}} K_{jp} D_{js} D_{jt}}{np^{1/2}} \right\|_{3a} \\
& \leq \left\| \frac{\sum_{s=1}^{t-\mathfrak{p}} \sum_{j=1}^{p} K_{jp} D_{js} \widetilde{D}_{jt}}{np^{1/2}} \right\|_{3a} + \frac{1}{np^{1/2}} \sum_{j=1}^{p} \left| K_{jp} \right| \left\| \sum_{s=1}^{t-\mathfrak{p}} D_{js} \right\|_{6a} \left\| D_{jt} - \widetilde{D}_{jt} \right\|_{6a} \\
& \leq \frac{C}{np^{1/2}} \left( \sum_{s=1}^{t-\mathfrak{p}} \left\| \sum_{j=1}^{p} K_{jp} D_{js} \widetilde{D}_{jt} \right\|_{3a}^{2} \right)^{1/2} + C \frac{\Theta_{6a} \left( \mathfrak{p} - p \right)}{p^{1/2}} \\
& \leq \frac{C}{np^{1/2}} \left( \left| t - \mathfrak{p} \right| p \right)^{1/2} + C \frac{\Theta_{6a} \left( \mathfrak{p} - p \right)}{p^{1/2}} \leq C \left( \frac{1}{n^{1/2}} + \frac{\Theta_{6a} \left( \mathfrak{p} - p \right)}{p^{1/2}} \right).
\end{aligned}
$$

Hence substituting gives

$$
\left\| \frac{\sum_{s=1}^{t-1} \left( \sum_{j=1}^{p} K_{jp} D_{js} D_{jt} \right)}{np^{1/2}} \right\|_{3a} \leq C \left( \frac{1}{n^{1/2}} + \frac{p^{1/2}\mathfrak{p}^{1/2}}{n} + \frac{\Theta_{6a} \left( \mathfrak{p} - p \right)}{p^{1/2}} \right). \tag{B.26}
$$

For the first item in (B.25), (B.23) gives a bound $C/n^{1/2}$. For the second item in (B.25), conditional Gaussianity of the $\left\{\sum_{j=1}^{p} K_{jp} D_{jt}\eta_{js}\right\}$ and Lemma B.3 give

$$\begin{aligned}
&\left\|\frac{\sum_{s=t+1}^{n}\left(\sum_{j=1}^{p} K_{jp} D_{jt}\eta_{js}\right)}{np^{1/2}}\right\|_{3a} \\
&\leq \frac{C}{np^{1/2}}\left\|\left\{\sum_{s=t+1}^{n}\left(\sum_{j=1}^{p} K_{jp}^{2} D_{jt}^{2}\right)\right\}^{1/2}\right\|_{3a} \leq \frac{C}{np^{1/2}}\left\|\sum_{s=t+1}^{n}\left(\sum_{j=1}^{p} K_{jp}^{2} D_{jt}^{2}\right)\right\|_{3a/2}^{1/2} \\
&\leq \frac{C}{np^{1/2}}\left(\sum_{s=t+1}^{n}\sum_{j=1}^{p} K_{jp}^{2}\left\|D_{jt}\right\|_{3a}^{2}\right)^{1/2} \leq \frac{C}{np^{1/2}}\left((n-t)\,p\right)^{1/2} \leq \frac{C}{n^{1/2}}.
\end{aligned}$$

Substituting the two last bounds in (B.25) and (B.26) in (B.24) shows that

$$\max\left\{\left\|\check{s}_{pt}^{(1)}\left(x;D_{t}\right)\right\|_{3a},\left\|\check{s}_{pt}^{(1)}\left(x;\eta_{t}\right)\right\|_{3a}\right\} \leq C\left(\frac{1}{n^{1/2}}+\frac{p^{1/2}\mathfrak{p}^{1/2}}{n}+\frac{\Theta_{6a}\left(\mathfrak{p}-p\right)}{p^{1/2}}\right). \tag{B.27}$$

Observe that $\Theta_{6a}\left(\mathfrak{p}-p\right) \leq C\left(\mathfrak{p}-p\right)^{-11/2}$ by Assumption R. Consider now

$$\mathfrak{p} = \max\left(2p, \left(\frac{n}{p}\right)^{\frac{1}{6}}\right) \geq 2p,$$

which is such that, since $p \in [1, \overline{p}_{n}]$ with $\overline{p}_{n} = O\left(n^{1/2}\right)$,

$$\begin{aligned}
&\text{If } \left(\frac{n}{p}\right)^{\frac{1}{6}} \geq 2p,\ \frac{\left(\mathfrak{p}-p\right)^{-11/2}}{p^{1/2}} \asymp \frac{p^{1/2}\mathfrak{p}^{1/2}}{n} \leq \frac{\mathfrak{p}}{n} \leq \frac{1}{n^{5/6}} \leq \frac{1}{n^{1/2}}, \\
&\text{If } \left(\frac{n}{p}\right)^{\frac{1}{6}} < 2p \Leftrightarrow \left(\frac{n}{2^{6}}\right)^{\frac{1}{7}} < p,\ \frac{\Theta_{6a}\left(\mathfrak{p}-p\right)}{p^{1/2}} \leq Cp^{-6} \leq \frac{C}{n^{1/2}},\ \frac{p^{1/2}\mathfrak{p}^{1/2}}{n} \leq \frac{\overline{p}_{n}}{n} \leq \frac{C}{n^{1/2}}.
\end{aligned}$$

Hence (B.27) gives (B.22). □

Let $I(\cdot)$ be a three times differentiable real function and define for $\mathcal{M}_{t}(\eta)$ as in (B.17),

$$\mathcal{I}_{t}(\eta) = \mathcal{I}_{tn}(\eta) = I\left(\mathcal{M}_{t}(\eta)\right), \quad \mathcal{I}_{t}(x;\eta) = \mathcal{I}(x\eta), \quad \mathcal{I}_{t}^{(j)}(x;\eta) = \frac{d_{t}^{j}\mathcal{I}(x;\eta)}{d^{j}x}, \quad j = 1,2.$$

Observe that $I(\mathcal{M}) = I\left(\mathcal{M}_{n}\left(D_{n}\right)\right) = \mathcal{I}_{n}\left(D_{n}\right)$, $\mathcal{I}_{t}\left(D_{t}\right) = \mathcal{I}_{t+1}\left(\eta_{t+1}\right)$, and that $I\left(\mathcal{M}_{1}\left(\eta_{1}\right)\right) = \mathcal{I}_{1}\left(\eta_{1}\right)$ is a function of the Gaussian vectors $\eta_{1},\ldots,\eta_{n}$ only.

**Lemma B.5.** *Let $\mathcal{M}$ and $\mathcal{M}_1(\eta_1)$ be as in (B.16) and (B.17). Consider a real function $I(\cdot)$ which may depend on $n$ and three times continuously differentiable with $\max_{j=1,2,3}\sup_x \left|I^{(j)}(x)\right| \leq C$. Then under Assumptions P, R and if $e = O\left(\overline{p}_n^{1/(2a)}\right)$,*

$$\left|\mathbb{E}\left[I(\mathcal{M}) - I(\mathcal{M}_1(\eta_1))\right]\right| \leq C\left(\frac{\overline{p}_n^{1+3/a}}{n^{1/2}} + \frac{1}{\overline{p}_n^{1-1/a}}\right).$$

**Proof of Lemma B.5.** The proof of the Lemma works by changing $D_n$ into $\eta_n$, $D_{n-1}$ into $\eta_{n-1}$ and so on, the so called Lindeberg technique described in Pollard (2002, p.179). This amounts to decompose $I(\mathcal{M}) - I(\mathcal{M}_n(\eta_n))$ into the following sum of differences,

$$\begin{aligned}
& I(\mathcal{M}) - I(\mathcal{M}_n(\eta_n)) \\
& = \mathcal{I}_n(D_n) - \mathcal{I}_{n-1}(D_{n-1}) + \mathcal{I}_{n-1}(D_{n-1}) - \mathcal{I}_{n-2}(D_{n-2}) + \cdots + \mathcal{I}_1(D_1) - \mathcal{I}_1(\eta_1) \\
& = \mathcal{I}_n(D_n) - \mathcal{I}_n(\eta_n) + \mathcal{I}_{n-1}(D_{n-1}) - \mathcal{I}_{n-1}(\eta_{n-1}) + \cdots + \mathcal{I}_1(D_1) - \mathcal{I}_1(\eta_1).
\end{aligned}$$

Since $\mathcal{I}_t(\eta) = \mathcal{I}_t(1;\eta)$ and $\mathcal{I}_t(0;\eta) = \mathcal{I}_t(0)$, a third-order Taylor expansion around $\eta = 0$ with integral remainder gives

$$\begin{aligned}
& [\mathcal{I}_t(D_t) - \mathcal{I}_t(\eta_t)] = \mathbb{E}\left[\mathcal{I}_t^{(1)}(0;D_t) - \mathcal{I}_t^{(1)}(0;\eta_t)\right] \\
& + \frac{1}{2}\mathbb{E}\left[\mathcal{I}_t^{(2)}(0;D_t) - \mathcal{I}_t^{(2)}(0;\eta_t)\right] + \frac{1}{2}\int_0^1 (1-x)^2 \mathbb{E}\left[\mathcal{I}_t^{(3)}(x;D_t) - \mathcal{I}_t^{(3)}(x;\eta_t)\right] dx.
\end{aligned}$$

Since $\{D_t\}$ is a sequence of martingale difference, $\mathbb{E}\left[\mathcal{I}_t^{(1)}(0;D_t) - \mathcal{I}_t^{(1)}(0;\eta_t)\right] = 0$ due to the expression of $\mathcal{I}_t^{(1)}(0;\eta)$ given above. Hence

$$\left|\mathbb{E}\left[I(\mathcal{M})\right] - \mathbb{E}\left[I(\mathcal{M}_1(\eta_1))\right]\right| \leq \frac{1}{2}\left|\sum_{t=1}^{n}\mathbb{E}\left[\mathcal{I}_t^{(2)}(0;D_t) - \mathcal{I}_t^{(2)}(0;\eta_t)\right]\right| \tag{B.28}$$

$$+ \frac{1}{2}\int_0^1 (1-x)^2 \left\{\sum_{t=1}^{n}\left|\mathbb{E}\left[\mathcal{I}_t^{(3)}(x;D_t) - \mathcal{I}_t^{(3)}(x;\eta_t)\right]\right|\right\} dx. \tag{B.29}$$

We now compute the differentials $\mathcal{I}_t^{(j)}(x;\eta)$, $j=1,2,3$. We have

$$\mathcal{I}_t^{(1)}(x;\eta) = I'(\mathcal{M}_t(x;\eta))\,\mathcal{M}_t^{(1)}(x;\eta),$$

$$\mathcal{I}_t^{(2)}(x;\eta) = I''(\mathcal{M}_t(x;\eta))\left(\mathcal{M}_t^{(1)}(x;\eta)\right)^2 + I'(\mathcal{M}_t(x;\eta))\,\mathcal{M}_t^{(2)}(x;\eta),$$

$$\begin{aligned}\mathcal{I}_t^{(3)}(x;\eta) &= I'''(\mathcal{M}_t(x;\eta))\left(\mathcal{M}_t^{(1)}(x;\eta)\right)^3 + 3I''(\mathcal{M}_t(x;\eta))\,\mathcal{M}_t^{(1)}(x;\eta)\,\mathcal{M}_t^{(2)}(x;\eta)\\ &\quad + I'(\mathcal{M}_t(x;\eta))\,\mathcal{M}_t^{(3)}(x;\eta).\end{aligned}$$

We compute the differentials of $\mathcal{M}_t$. We have

$$\begin{aligned}\mathcal{M}_t^{(1)}(x;\eta) &= \left(\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{e}(x;\eta)\right)^{1/e-1}\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{e-1}(x;\eta)\,\Sigma_{pt}^{(1)}(x;\eta)\\ &= \mathcal{M}_t^{1-e}(x;\eta)\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{e-1}(x;\eta)\,\Sigma_{pt}^{(1)}(x;\eta),\\ \mathcal{M}_t^{(2)}(x;\eta) &= \mathcal{M}_{1t}^{(2)}(x;\eta)+\mathcal{M}_{2t}^{(2)}(x;\eta)+\mathcal{M}_{3t}^{(2)}(x;\eta),\\ \mathcal{M}^{(3)}(x;\eta) &= \mathcal{M}_{1t}^{(3)}(x;\eta)+\cdots+\mathcal{M}_{6t}^{(3)}(x;\eta),\end{aligned}$$

where, dropping the variables $x$, $\eta$ for notational convenience

$$\mathcal{M}_{1t}^{(2)} = \left(\frac{1}{e} - 1\right) \mathcal{M}_t^{1-2e} \left(\sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(1)}\right)^2,$$

$$\mathcal{M}_{2t}^{(2)} = \mathcal{M}_t^{1-e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(2)},$$

$$\mathcal{M}_{3t}^{(2)} = (e-1) \mathcal{M}_t^{1-e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-2} \left(\Sigma_{pt}^{(1)}\right)^2,$$

$$\mathcal{M}_{1t}^{(3)} = \left(\frac{1}{e} - 1\right) \left(\frac{1}{e} - 2\right) \mathcal{M}_t^{1-3e} \left(\sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(1)}\right)^3,$$

$$\mathcal{M}_{2t}^{(3)} = 3 \left(\frac{1}{e} - 1\right) \mathcal{M}_t^{1-2e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(1)} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(2)},$$

$$\mathcal{M}_{3t}^{(3)} = 3 \left(\frac{1}{e} - 1\right) (e-1) \mathcal{M}_t^{1-2e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(1)} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-2} \left(\Sigma_{pt}^{(1)}\right)^2,$$

$$\mathcal{M}_{4t}^{(3)} = (3e-1) \mathcal{M}_t^{1-e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-2} \Sigma_{pt}^{(2)} \Sigma_{pt}^{(1)},$$

$$\mathcal{M}_{5t}^{(3)} = (e-1)(e-2) \mathcal{M}_t^{1-e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-2} \left(\Sigma_{pt}^{(1)}\right)^3,$$

$$\mathcal{M}_{6t}^{(3)} = \mathcal{M}_t^{1-e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(3)}.$$

**The third-order item(B.29)**. Since

$$\begin{aligned}
&\frac{1}{2} \int_0^1 (1-x)^2 \left\{ \sum_{t=1}^{n} \left| \mathbb{E}\left[ \mathcal{I}_t^{(3)}(x; D_t) - \mathcal{I}_t^{(3)}(x; \eta_t) \right] \right| \right\} dx \\
&\leq \frac{1}{2} \int_0^1 (1-x)^2 \left\{ \sum_{t=1}^{n} \left( \left| \mathbb{E}\left[ \mathcal{I}_t^{(3)}(x; D_t) \right] \right| + \left| \mathbb{E}\left[ \mathcal{I}_t^{(3)}(x; \eta_t) \right] \right| \right) \right\} dx,
\end{aligned}$$

it is sufficient to bound $\sum_{t=1}^{n}\left|\mathbb{E}\left[\mathcal{I}_t^{(3)}(x)\right]\right|$ independently of $x$ where $\mathcal{I}_t^{(3)}(x)$ stands for $\mathcal{I}_t^{(3)}(x;\eta_t)$ or $\mathcal{I}_t^{(3)}(x;D_t)$. We have, dropping dependence w.r.t. to $x$ for ease of notation,

$$\begin{aligned}\sum_{t=1}^{n}\left|\mathbb{E}\left[\mathcal{I}_t^{(3)}\right]\right| \leq C\sum_{t=1}^{n}\left\{\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\right|^3\right]+\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{1t}^{(2)}\right|\right]+\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{2t}^{(2)}\right|\right]\right\}\\ +C\sum_{t=1}^{n}\left\{\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{3t}^{(2)}\right|\right]+\sum_{j=1}^{6}\mathbb{E}\left[\left|\mathcal{M}_{jt}^{(3)}\right|\right]\right\}.\end{aligned}$$

We now study the ten items above.

**(1)** $\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\right|^3\right]$. We have for $a$, $\overline{a}\geq 1$ with $1/a = 1-1/\overline{a}$,

$$\begin{aligned}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\right|^3\right] &= \mathbb{E}\left[\left|\mathcal{M}_t^{1-e}\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{e-1}\Sigma_{pt}^{(1)}\right|^3\right]\\ &\leq \sum_{p_1,p_2,p_3=1}^{\overline{p}_n}\mathbb{E}\left[\left|\mathcal{M}_t^{3(1-e)}\Sigma_{p_1t}^{e-1}\Sigma_{p_2t}^{e-1}\Sigma_{p_3t}^{e-1}\Sigma_{p_1t}^{(1)}\Sigma_{p_2t}^{(1)}\Sigma_{p_3t}^{(1)}\right|\right]\\ &\leq \max_{p,t}\left\|\Sigma_{pt}^{(1)}\right\|_{3a}^3\sum_{p_1,p_2,p_3=1}^{\overline{p}_n}\mathbb{E}^{1/\overline{a}}\left[\left|\mathcal{M}_t^{3(1-e)}\Sigma_{p_1t}^{e-1}\Sigma_{p_2t}^{e-1}\Sigma_{p_3t}^{e-1}\right|^{\overline{a}}\right]\\ &\leq \frac{C}{n^{3/2}}\sum_{p_1,p_2,p_3=1}^{\overline{p}_n}\mathbb{E}^{1/\overline{a}}\left[\left|\mathcal{M}_t^{3(1-e)}\Sigma_{p_1t}^{e-1}\Sigma_{p_2t}^{e-1}\Sigma_{p_3t}^{e-1}\right|^{\overline{a}}\right],\end{aligned}$$

by (B.18) for all $x \in [0,1]$. Now, since $t \mapsto t^{1/\overline{a}}$, $t \mapsto t^{1-1/e}$ are concave and $\sum_{p=1}^{\overline{p}_n} t_p^{\overline{a}} \leq \left(\sum_{p=1}^{\overline{p}_n} t_p\right)^{\overline{a}}$, the definition of $\mathcal{M}_t$ gives

$$\begin{aligned}
&\sum_{p_1,p_2,p_3=1}^{\overline{p}_n} \mathbb{E}^{1/\overline{a}}\left[\left|\mathcal{M}_t^{3(1-e)}\Sigma_{p_1t}^{e-1}\Sigma_{p_2t}^{e-1}\Sigma_{p_3t}^{e-1}\right|^{\overline{a}}\right] \\
&= \overline{p}_n^3 \times \frac{1}{\overline{p}_n^3}\sum_{p_1,p_2,p_3=1}^{\overline{p}_n} \mathbb{E}^{1/\overline{a}}\left[\left|\mathcal{M}_t^{3(1-e)}\Sigma_{p_1t}^{e-1}\Sigma_{p_2t}^{e-1}\Sigma_{p_3t}^{e-1}\right|^{\overline{a}}\right] \\
&\leq \overline{p}_n^3\left(\frac{1}{\overline{p}_n^3}\mathbb{E}\left[\sum_{p_1,p_2,p_3=1}^{\overline{p}_n}\mathcal{M}_t^{3\overline{a}(1-e)}\Sigma_{p_1t}^{\overline{a}e(1-1/e)}\Sigma_{p_2t}^{\overline{a}e(1-1/e)}\Sigma_{p_3t}^{\overline{a}e(1-1/e)}\right]\right)^{1/\overline{a}} \\
&= \overline{p}_n^3\left(\mathbb{E}\left[\left(\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{e}\right)^{-3\overline{a}(1-1/e)}\left(\frac{1}{\overline{p}_n}\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{\overline{a}e(1-1/e)}\right)^3\right]\right)^{1/\overline{a}} \\
&\leq \overline{p}_n^3\left(\mathbb{E}\left[\left(\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{\overline{a}e}\right)^{-3(1-1/e)}\left(\frac{1}{\overline{p}_n}\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{\overline{a}e}\right)^{3(1-1/e)}\right]\right)^{1/\overline{a}} \\
&\leq \overline{p}_n^{3(1-1/\overline{a})+3/(e\overline{a})} \leq C\overline{p}_n^{3/a},
\end{aligned}$$

uniformly w.r.t. to $t$ since $(\ln \overline{p}_n)/e = o(1)$. Hence for all $x \in [0,1]$

$$\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\right|^3\right] \leq C\frac{\overline{p}_n^{3/a}}{n^{1/2}}. \tag{B.30}$$

**(2)** $\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{1t}^{(2)}\right|\right]$. We have, since $\mathcal{M}_t \geq 1$,

$$\begin{aligned}
\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\right|\left|\mathcal{M}_{1t}^{(2)}\right|\right] &\leq C\mathbb{E}\left[\mathcal{M}_t^{2-3e}\left|\sum_{p=2}^{\overline{p}_n}\Sigma_{pt}^{e-1}\Sigma_{pt}^{(1)}\right|^3\right] \leq C\mathbb{E}\left[\mathcal{M}_t^{3-3e}\left|\sum_{p=2}^{\overline{p}_n}\Sigma_{pt}^{e-1}\Sigma_{pt}^{(1)}\right|^3\right] \\
&\leq C\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\right|^3\right],
\end{aligned}$$

for all $t$, such that $\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\right|^2\left|\mathcal{M}_{1t}^{(2)}\right|\right] \leq C\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\right|^3\right]$. Hence a bound similar to (B.30) holds.

**(3)** $\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{2t}^{(2)}\right|\right]$. Let $\overline{a}>1$ be such that $1/\overline{a}=1-1/a$. Arguing as for **(1)** with (B.18) and (B.19),

$$\begin{aligned}
\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{1t}^{(2)}\right|\right] &\leq C\sum_{p_1,p_2=1}^{\overline{p}_n}\mathbb{E}\left[\mathcal{M}_t^{2(1-e)}\left|\Sigma_{p_1t}^{e-1}\Sigma_{p_2t}^{e-1}\Sigma_{p_1t}^{(1)}\Sigma_{p_2t}^{(2)}\right|\right] \\
&\leq C\max_{p,t}\left\{\left\|\Sigma_{pt}^{(1)}\right\|_{3a}\left\|\Sigma_{pt}^{(2)}\right\|_{3a/2}\right\}\sum_{p_1,p_2=1}^{\overline{p}_n}\mathbb{E}^{1/\overline{a}}\left[\left|\mathcal{M}_t^{2(1-e)}\Sigma_{p_1t}^{e-1}\Sigma_{p_2t}^{e-1}\right|^{\overline{a}}\right] \\
&\leq C\frac{\overline{p}_n^{1/2}}{n^{3/2}}\times\overline{p}_n^2\times\mathbb{E}^{1/\overline{a}}\left[\left(\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{e}\right)^{-2\overline{a}(1-1/e)}\left(\frac{1}{\overline{p}_n}\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{e\overline{a}(1-1/e)}\right)^2\right] \\
&\leq C\frac{\overline{p}_n^{1/2}}{n^{3/2}}\times\overline{p}_n^2\times\mathbb{E}^{1/\overline{a}}\left[\left(\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{e\overline{a}(1-1/e)}\right)^{-2}\left(\frac{1}{\overline{p}_n}\sum_{p=1}^{\overline{p}_n}\Sigma_{pt}^{e\overline{a}(1-1/e)}\right)^2\right] \\
&= C\frac{\overline{p}_n^{1/2}}{n^{3/2}}\times\overline{p}_n^2\times\overline{p}_n^{-2/\overline{a}} = C\frac{\overline{p}_n^{\frac{1}{2}(1+4/a)}}{n^{3/2}}.
\end{aligned}$$

Hence, uniformly w.r.t. $x\in[0,1]$,

$$\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{2t}^{(2)}\right|\right]\leq C\frac{\overline{p}_n^{\frac{1}{2}(1+4/a)}}{n^{1/2}}. \tag{B.31}$$

**(4)** $\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{3t}^{(2)}\right|\right]$. Proceeding as **(1)** and **(3)** gives, since $\inf_{p,t}\Sigma_{pt}\geq 1$,

$$\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{3t}^{(2)}\right|\right]\leq Ce\sum_{p_1,p_2=1}^{\overline{p}_n}\mathbb{E}\left[\mathcal{M}_t^{2(1-e)}\left|\Sigma_{p_1t}^{e-1}\Sigma_{p_2t}^{e-1}\Sigma_{p_1t}^{(1)}\left(\Sigma_{p_2t}^{(1)}\right)^2\right|\right]\leq C\frac{e\overline{p}_n^{2/a}}{n^{3/2}}\leq C\frac{\overline{p}_n^{3/a}}{n^{3/2}},$$

provided $e=O(\overline{p}_n^{1/a})$. Hence $\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_t^{(1)}\mathcal{M}_{3t}^{(2)}\right|\right]$ can be bounded as in (B.30).

**(5)** $\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_{1t}^{(3)}\right|\right]$ can be bounded as in (B.30) since $\mathcal{M}_t\geq 1$ gives $\mathbb{E}\left[\left|\mathcal{M}_{1t}^{(3)}\right|\right]\leq C\mathbb{E}\left[\mathcal{M}_t^{3(1-e)}\left|\sum_{p=2}^{\overline{p}_n}\Sigma_{pt}^{e-1}\Sigma_{pt}^{(1)}\right|^3\right]$.

**(6)** $\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_{2t}^{(3)}\right|\right]$. Arguing as in **(3)** gives that $\sum_{t=1}^{n}\mathbb{E}\left[\left|\mathcal{M}_{2t}^{(3)}\right|\right]$ can be bounded as in (B.31).

**(7)** $\sum_{t=1}^{n} \mathbb{E}\left[\left|\mathcal{M}_{3t}^{(3)}\right|\right]$. Arguing as in **(4)** shows that this item is negligible compared to (B.30).

**(8)** $\sum_{t=1}^{n} \mathbb{E}\left[\left|\mathcal{M}_{4t}^{(3)}\right|\right]$. Let $\overline{a} > 1$ be such that $1/\overline{a} = 1 - 1/a$. We have, since $\inf_{p,t} \Sigma_{pt} \geq 1$,

$$\mathbb{E}\left[\left|\mathcal{M}_{4t}^{(3)}\right|\right] \leq Ce\mathbb{E}\left[\mathcal{M}_t^{1-e} \sum_{p=1}^{\overline{p}_n} \left|\Sigma_{pt}^{e-2}\Sigma_{pt}^{(2)}\Sigma_{pt}^{(1)}\right|\right] \leq Ce \sum_{p=p_o}^{\overline{p}_n} \mathbb{E}^{1/\overline{a}}\left[\left(\mathcal{M}_t^{1-e}\Sigma_{pt}^{e-1}\right)^{\overline{a}}\right] \left\|\Sigma_{pt}^{(2)}\right\|_{3a/2} \left\|\Sigma_{pt}^{(1)}\right\|_{3a}$$

$$\leq C\frac{e\overline{p}_n^{1/2}\overline{p}_n^{1-1/\overline{a}}}{n^{3/2}} \leq C\frac{\overline{p}_n^{\frac{1}{2}(1+4/a)}}{n^{3/2}},$$

provided $e = O\left(\overline{p}_n^{1/a}\right)$. This gives a bound similar to (B.31) for $\sum_{t=1}^{n} \mathbb{E}\left[\left|\mathcal{M}_{4t}^{(3)}\right|\right]$.

**(9)** $\sum_{t=1}^{n} \mathbb{E}\left[\left|\mathcal{M}_{5t}^{(3)}\right|\right]$ can be bounded as in (B.30) provided $e = O(\overline{p}_n^{1/(2a)})$.

**(10)** $\sum_{t=1}^{n} \mathbb{E}\left[\left|\mathcal{M}_{6t}^{(3)}\right|\right]$ can be bounded as in (B.31).

Hence, collecting the dominant bounds (B.30) and (B.31) in **(1)**-**(10)** gives

$$\frac{1}{2}\int_0^1 (1-x)^2 \left\{\sum_{t=1}^{n} \left|\mathbb{E}\left[\mathcal{I}_t^{(3)}(x; D_t) - \mathcal{I}_t^{(3)}(x; \eta_t)\right]\right|\right\} dx \leq C\frac{\overline{p}_n^{\frac{3}{a}} + \overline{p}_n^{\frac{1}{2}(1+4/a)}}{n^{1/2}} \leq C\left(\frac{\overline{p}_n^{1+\frac{4}{a}}}{n}\right)^{\frac{1}{2}}. \tag{B.32}$$

**The second-order term (B.28)**. Note that $\mathcal{I}_t^{(2)}(0;\eta) = \eta' A_t \eta$ where $A_t$ depends upon $D_1, \ldots, D_{t-1}$ and $\eta_{t+1}, \ldots, \eta_n$. In the standard Lindeberg method, $\{D_t, t \in [1,n]\}$ and $\{\eta_t, t \in [1,n]\}$ are both independent variables with identical mean and variance, so that the second order term, which writes as a sum of items $\mathbb{E}\left[D_t' A_t D_t\right] - \mathbb{E}\left[\eta_t' A_t \eta_t\right]$, is equal to 0 in this simpler case. However this does not hold in our case. In this step, the second order term is dealt with by removing from $\mathcal{I}_t^{(2)}(0;\eta)$ a block $\sum_{j=1}^{p} K_{jp} \sum_{s=t-\ell}^{t-1} D_{js}$ and by changing the $D_{jt}$ into $D_{jt}^{t-\ell+1} = \mathbb{E}\left[D_{jt} \left| e_t, \ldots, e_{t-\ell+1}\right.\right]$.

Observe that $\mathcal{I}_t^{(2)}(0;\eta) = \mathcal{I}_{1t}^{(2)}(0;\eta) + \mathcal{I}_{2t}^{(2)}(0;\eta) + \mathcal{I}_{3t}^{(2)}(0;\eta) + \mathcal{I}_{4t}^{(2)}(0;\eta)$ with, dropping the dependence upon 0 and $\eta$,

$$\mathcal{I}_{1t}^{(2)} = \left(\frac{1}{e} - 1\right) I_{tn}^{(1)} \mathcal{M}_t^{1-2e} \left(\sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(1)}\right)^2, \quad I_{tn}^{(1)} = I'\left(\mathcal{M}_t\right),$$

$$\mathcal{I}_{2t}^{(2)} = I_{tn}^{(1)} \mathcal{M}_t^{1-e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(2)}, \quad \mathcal{I}_{3t}^{(2)} = (e-1)\, I_{tn}^{(1)} \mathcal{M}_t^{1-e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \left(\Sigma_{pt}^{(1)}\right)^2,$$

$$\mathcal{I}_{4t}^{(2)} = I''\left(\mathcal{M}_t\right) \left(\mathcal{M}_t^{1-e} \sum_{p=1}^{\overline{p}_n} \Sigma_{pt}^{e-1} \Sigma_{pt}^{(1)}\right)^2.$$

Observe $\mathcal{M}_t\left(0; D_t\right) = \mathcal{M}_t\left(0; \eta_t\right)$ and $\Sigma_{pt}\left(0; D_t\right) = \Sigma_{pt}\left(0; \eta_t\right)$ and that these quantities do not depend upon $\eta_t$ or $D_t$. We shall first focus on $\mathcal{I}_{1t}^{(2)}$. Let $\ell \geq 2\overline{p}_n$ be an integer number. Define, for $y \in [0, 1]$,

$$\mathfrak{S}_{pt}\left(y; \eta\right) = \frac{2 \sum_{j=1}^{p} K_{jp} \left(\sum_{s=j+1}^{t-\ell-1} D_{js} + y \sum_{s=t-\ell}^{t-1} D_{js} + \sum_{s=t+1}^{n} \eta_{js}\right) \eta_j}{n\sigma^4 V_\Delta\left(p\right)},$$

$$\mathfrak{S}_{pt}\left(y\right) = \mathfrak{S}_{pt}\left(y; yD_t + (1-y)\, D_t^{t-\ell+1}\right),$$

$$\mathfrak{T}_{pt}\left(y; \eta\right) = \check{s}_{pt}^{(2)}(y; \eta) = \frac{2 \sum_{j=1}^{p} K_{jp} \eta_j^2}{n\sigma^4 V_\Delta\left(p\right)}, \quad \mathfrak{T}_{pt}\left(y\right) = \mathfrak{T}_{pt}\left(y; yD_t + (1-y)\, D_t^{t-\ell+1}\right),$$

which are such that $\mathfrak{S}_{pt}(1;\eta) = \check{s}^{(1)}_{pt}(0;\eta)$, $\mathfrak{S}_{pt}(1) = \check{s}^{(1)}_{pt}(0;D_t)$, $\mathfrak{T}_{pt}(1) = \check{s}^{(2)}_{pt}(0;D_t)$. Define also

$$
\begin{aligned}
\mathbf{M}_{jt}(y) &= \sum_{s=j+1}^{t-\ell-1} D_{js} + y\sum_{s=t-\ell}^{t-1} D_{js} + \sum_{s=t+1}^{n} \eta_{js}, \quad \mathbf{R}_{jt}(y) = \frac{\mathbf{M}_{jt}(y)}{n}, \\
\mathbf{s}_{pt}(y) &= \frac{n\sum_{j=1}^{p} K_{jp}\mathbf{R}^2_{jt}(y) - \sigma^4 E_\Delta(p)}{\sigma^4 V_\Delta(p)}, \quad \mathbf{\Sigma}_{pt}(y) = f(\mathbf{s}_{pt}(y)), \\
\widetilde{\Sigma}^{(1)}_{pt}(y;\eta) &= f^{(1)}(\mathbf{s}_{pt}(y))\,\mathfrak{S}_{pt}(y;\eta), \\
\widetilde{\Sigma}^{(2)}_{pt}(y;\eta) &= f^{(1)}(\mathbf{s}_{pt}(y))\,\mathfrak{T}_{pt}(y;\eta) + f^{(2)}(\mathbf{s}_{pt}(y))\,(\mathfrak{S}_{pt}(y;\eta))^2, \\
\widetilde{\Sigma}^{(1)}_{pt}(y) &= \widetilde{\Sigma}^{(1)}_{pt}\left(y; yD_t + (1-y)\,D_t^{t-\ell+1}\right), \\
\widetilde{\Sigma}^{(2)}_{pt}(y;\eta) &= \widetilde{\Sigma}^{(2)}_{pt}\left(y; yD_t + (1-y)\,D_t^{t-\ell+1}\right), \\
\mathfrak{M}_t(y) &= \left(\sum_{p=1}^{\overline{p}_n} \mathbf{\Sigma}^e_{pt}(y)\right)^{\frac{1}{e}}, \quad \mathfrak{I}^{(1)}_{tn}(y) = I'(\mathfrak{M}_t(y)),
\end{aligned}
$$

and the counterpart of $\mathcal{I}^{(2)}_{1t}(0;\eta_t)$ and $\mathcal{I}^{(2)}_{1t}(0;D_t)$ as

$$
\begin{aligned}
\mathfrak{I}_t(y;\eta) &= \left(\frac{1}{e}-1\right)\mathfrak{I}^{(1)}_{tn}(y)\,\mathfrak{M}^{1-2e}_t(y)\left(\sum_{p=1}^{\overline{p}_n} \mathbf{\Sigma}^{e-1}_{pt}(y)\,\widetilde{\Sigma}^{(1)}_{pt}(y;\eta)\right)^2, \\
\mathfrak{I}_t(y) &= \mathfrak{I}_t\left(y; yD_t + (1-y)\,D_t^{t-\ell+1}\right).
\end{aligned}
$$

Observe that $\mathcal{I}^{(2)}_{1t}(0;\eta_t) = \mathfrak{I}_t(1;\eta_t)$ and $\mathcal{I}^{(2)}_{1t}(0;D_t) = \mathfrak{I}_t(1)$. Hence $\mathbb{E}\left[\mathcal{I}^{(2)}_{1t}(0;D_t) - \mathcal{I}^{(2)}_{1t}(0;\eta_t)\right] = \mathbb{E}\left[\mathfrak{I}_t(1) - \mathfrak{I}_t(1;\eta_t)\right]$ and

$$
\begin{aligned}
\mathbb{E}\left[\mathcal{I}^{(2)}_{1t}(0;D_t) - \mathcal{I}^{(2)}_{1t}(0;\eta_t)\right] &= \mathbb{E}\left[\mathfrak{I}_t(0) - \mathfrak{I}_t(0;\eta_t)\right] && \text{(B.33)} \\
&\quad + \int_0^1 \mathbb{E}\left[\mathfrak{I}^{(1)}_t(y) - \mathfrak{I}^{(1)}_t(y;\eta_t)\right] dy, && \text{(B.34)}
\end{aligned}
$$

where $\mathfrak{I}^{(1)}_t(y) = d\mathfrak{I}_t(y)/dy$ and $\mathfrak{I}^{(1)}_t(y;\eta_t) = d\mathfrak{I}_t(y;\eta_t)/dy$.

We first consider the integral item $\int_0^1 \left|\mathbb{E}\left[\mathfrak{I}_t^{(1)}(y)\right]\right| dy$ from (B.34) and first compute $\mathfrak{I}_{1t}^{(1)}(y)$. Define

$$\mathfrak{S}_{pt}^{(1)}(y) = \frac{d\mathfrak{S}_{pt}(y)}{dy} = \frac{2\sum_{j=1}^{p} K_{jp}\left(\sum_{s=t-\ell}^{t-1} D_{js}\right)\left(yD_{jt} + (1-y)D_{jt}^{t-\ell+1}\right)}{n\sigma^4 V_\Delta(p)}$$
$$+ \frac{2\sum_{j=1}^{p} K_{jp}\left(\sum_{s=j+1}^{t-\ell-1} D_{js} + y\sum_{s=t-\ell}^{t-1} D_{js} + \sum_{s=t+1}^{n} \eta_{js}\right)\left(D_{jt}^{t-\ell+1} - D_{jt}\right)}{n\sigma^4 V_\Delta(p)},$$

$$\mathfrak{T}_{pt}^{(1)}(y) = \frac{d\mathfrak{T}_{pt}(y)}{dy} = \frac{4\sum_{j=1}^{p} K_{jp}\left(yD_{jt} + (1-y)D_{jt}^{t-\ell+1}\right)\left(D_{jt} - D_{jt}^{t-\ell+1}\right)}{n\sigma^4 V_\Delta(p)},$$

$$\mathbf{s}_{pt}^{(1)}(y) = \frac{d\mathbf{s}_{pt}(y)}{dy} = \frac{2\sum_{j=1}^{p} K_{jp}\mathbf{M}_{jt}(y)\sum_{s=t-\ell}^{t-1} D_{js}}{n\sigma^4 V_\Delta(p)},$$

$$\widetilde{\Sigma}_{pt}^{(1,1)}(y) = \frac{d\widetilde{\Sigma}_{pt}^{(1)}(y)}{dy} = f^{(2)}(\mathbf{s}_{pt}(y))\,\mathbf{s}_{pt}^{(1)}(y)\,\mathfrak{S}_{pt}(y) + f^{(1)}(\mathbf{s}_{pt}(y))\,\mathfrak{S}_{pt}^{(1)}(y),$$

$$\widetilde{\Sigma}_{pt}^{(2,1)}(y) = \frac{d\widetilde{\Sigma}_{pt}^{(2)}(y)}{dy} = f^{(2)}(\mathbf{s}_{pt}(y))\,\mathbf{s}_{pt}^{(1)}(y)\,\mathfrak{T}_{pt}(y) + f^{(1)}(\mathbf{s}_{pt}(y))\,\mathfrak{T}_{pt}^{(1)}(y)$$
$$+ f^{(3)}(\mathbf{s}_{pt}(y))\,\mathbf{s}_{pt}^{(1)}(y)\left(\mathfrak{S}_{pt}(y)\right)^2 + 2f^{(2)}(\mathbf{s}_{pt}(y))\,\mathfrak{S}_{pt}(y)\,\mathfrak{S}_{pt}^{(1)}(y),$$

$$\mathfrak{I}_{tn}^{(2)}(y) = I''(\mathfrak{M}_t(y)),$$

and

$$\mathfrak{I}_{1t}^{(1)}(y) = \left(\frac{1}{e}-1\right)\mathfrak{I}_{tn}^{(2)}(y)\,\mathfrak{M}_t^{2-3e}(y)\left(\sum_{p=1}^{\bar{p}_n}\mathbf{\Sigma}_{pt}^{e-1}(y)\,\widetilde{\Sigma}_{pt}^{(1)}(y)\right)^2\sum_{p=1}^{\bar{p}_n}\mathbf{\Sigma}_{pt}^{e-1}(y)\,\mathbf{\Sigma}_{pt}^{(1)}(y),$$

$$\mathfrak{I}_{2t}^{(1)}(y) = \left(\frac{1}{e}-1\right)\left(\frac{1}{e}-2\right)\mathfrak{I}_{tn}^{(1)}(y)\,\mathfrak{M}_t^{1-3e}(y)\left(\sum_{p=1}^{\bar{p}_n}\mathbf{\Sigma}_{pt}^{e-1}(y)\,\widetilde{\Sigma}_{pt}^{(1)}(y)\right)^2\sum_{p=1}^{\bar{p}_n}\mathbf{\Sigma}_{pt}^{e-1}(y)\,\mathbf{\Sigma}_{pt}^{(1)}(y),$$

$$\mathfrak{I}_{3t}^{(1)}(y) = 2\left(\frac{1}{e}-1\right)(e-1)\,\mathfrak{I}_{tn}^{(1)}(y)\,\mathfrak{M}_t^{1-2e}(y)\left(\sum_{p=1}^{\bar{p}_n}\mathbf{\Sigma}_{pt}^{e-1}(y)\,\widetilde{\Sigma}_{pt}^{(1)}(y)\right)\left(\sum_{p=1}^{\bar{p}_n}\mathbf{\Sigma}_{pt}^{e-2}(y)\left(\mathbf{\Sigma}_{pt}^{(1)}(y)\right)^2\right),$$

$$\mathfrak{I}_{4t}^{(1)}(y) = 2\left(\frac{1}{e}-1\right)\mathfrak{I}_{tn}^{(1)}(y)\,\mathfrak{M}_t^{1-2e}(y)\left(\sum_{p=1}^{\bar{p}_n}\mathbf{\Sigma}_{pt}^{e-1}(y)\,\widetilde{\Sigma}_{pt}^{(1)}(y)\right)\left(\sum_{p=1}^{\bar{p}_n}\mathbf{\Sigma}_{pt}^{e-1}(y)\,\widetilde{\Sigma}_{pt}^{(1,1)}(y)\right).$$

To bound the moments of $\widetilde{\Sigma}_{pt}^{(1)}(y)$, $\widetilde{\Sigma}_{pt}^{(1,1)}(y)$ and $\mathbf{\Sigma}_{pt}^{(1)}(y)$, consider first $\left\|\mathfrak{S}_{pt}(y)\right\|_{3a}$, $\left\|\mathfrak{S}_{pt}^{(1)}(y)\right\|_{3a}$ and $\left\|\mathbf{s}_{pt}^{(1)}(y)\right\|_{3a}$. For $\left\|\mathfrak{S}_{pt}(y)\right\|_{3a}$ and $\left\|\mathfrak{S}_{pt}^{(1)}(y)\right\|_{3a}$, (B.18), the Burkholder inequality, (B.6) $\overline{p}_n = O\left(n^{1/2}\right)$, $2\overline{p}_n \leq \ell \leq 3\overline{p}_n$ and $\Theta_{6a}(\ell - \overline{p}_n) \leq C\overline{p}_n^{-1}$ give

$$
\begin{aligned}
&\left\|\mathfrak{S}_{pt}(y)\right\|_{3a} \\
&\leq \left\|\frac{2\sum_{j=1}^{p} K_{jp}\left(\sum_{s=j+1}^{t-\ell-1} D_{js} + y\sum_{s=t-\ell}^{t-1} D_{js} + \sum_{s=t+1}^{n}\eta_{js}\right) D_{jt}}{n\sigma^4 V_\Delta(p)}\right\|_{3a} \\
&+ 2\left|1-y\right|\sum_{j=1}^{p}\frac{\left|K_{jp}\right|}{n\sigma^4 V_\Delta(p)}\left\|\left(\sum_{s=j+1}^{t-\ell-1} D_{js} + y\sum_{s=t-\ell}^{t-1} D_{js} + \sum_{s=t+1}^{n}\eta_{js}\right)\right\|_{6a}\left\|D_{jt} - D_{jt}^{t-\ell+1}\right\|_{6a} \\
&\leq C\left(\frac{1}{n^{1/2}} + \frac{\overline{p}_n}{n} + \left(\frac{\overline{p}_n}{n}\right)^{1/2}\Theta_{6a}(\ell - \overline{p}_n)\right) \leq \frac{C}{n^{1/2}},
\end{aligned}
$$

$$
\begin{aligned}
&\left\|\mathfrak{S}_{pt}^{(1)}(y)\right\|_{3a} \\
&\leq \left\|\frac{2\sum_{j=1}^{p} K_{jp}\left(\sum_{s=t-\ell}^{t-1} D_{js}\right) D_{jt}}{n\sigma^4 V_\Delta(p)}\right\|_{3a} \\
&+ 2\left|1-y\right|\sum_{j=1}^{p}\frac{\left|K_{jp}\right|}{n\sigma^4 V_\Delta(p)}\left\|\sum_{s=t-\ell}^{t-1} D_{js}\right\|_{6a}\left\|D_{jt} - D_{jt}^{t-\ell+1}\right\|_{6a} \\
&+ 2\sum_{j=1}^{p}\frac{\left|K_{jp}\right|}{n\sigma^4 V_\Delta(p)}\left\|\sum_{s=j+1}^{t-\ell-1} D_{js} + y\sum_{s=t-\ell}^{t-1} D_{js} + \sum_{s=t+1}^{n}\eta_{js}\right\|_{6a}\left\|D_{jt} - D_{jt}^{t-\ell+1}\right\|_{6a} \\
&\leq C\left(\frac{\ell^{1/2}}{n} + \frac{\ell^{1/2}\overline{p}_n^{1/2}}{n}\Theta_{6a}(\ell - \overline{p}_n) + \left(\frac{\overline{p}_n}{n}\right)^{1/2}\Theta_{6a}(\ell - \overline{p}_n)\right) \\
&\leq C\left(\frac{\overline{p}_n^{1/2}}{n} + \frac{1}{\left(n\overline{p}_n\right)^{1/2}}\right),
\end{aligned}
$$

$$
\left\|\mathfrak{T}_{pt}(y)\right\|_{3a} \leq C\frac{\overline{p}_n^{1/2}}{n}, \quad \left\|\mathfrak{T}_{pt}^{(1)}(y)\right\|_{3a} \leq \frac{C}{n\overline{p}_n}.
$$

For $\left\| \mathbf{s}_{pt}^{(1)}(y) \right\|_{3a}$ (B.18), $\overline{p}_n = O\left(n^{1/2}\right)$ and the Burkholder inequality give

$$\begin{aligned}
&\left\| \mathbf{s}_{pt}^{(1)}(y) \right\|_{3a} \\
&\leq \left\| 2 \sum_{s_1=t-\ell}^{t-1} \sum_{j=1}^{p} \frac{K_{jp}}{n\sigma^4 V_\Delta(p)} \left( \sum_{s_2=j+1}^{t-\ell-1} D_{js_2} \right) D_{js_1} \right\|_{3a} + \left\| \frac{2 \sum_{j=1}^{p} K_{jp} \left( \sum_{s=t-\ell}^{t-1} D_{js} \right)^2}{n\sigma^4 V_\Delta(p)} \right\|_{3a} \\
&+ \left\| \frac{2 \sum_{j=1}^{p} K_{jp} \left( \sum_{s=t-\ell}^{t-1} D_{js} \right) \left( \sum_{s=t+1}^{n} \eta_{js} \right)}{n\sigma^4 V_\Delta(p)} \right\|_{3a} \\
&\leq C \left( \sum_{s_1=t-\ell}^{t-1} \left\| \sum_{j=1}^{p} \frac{K_{jp}}{n\sigma^4 V_\Delta(p)} \left( \sum_{s_2=j+1}^{t-\ell-1} D_{js_2} \right) D_{js_1} \right\|_{3a}^2 \right)^{1/2} + C \sum_{j=1}^{p} \frac{|K_{jp}|}{n\sigma^4 V_\Delta(p)} \left\| \sum_{s=t-\ell}^{t-1} D_{js} \right\|_{6a}^2 \\
&+ C \left\| \frac{\left( \sum_{j=1}^{p} K_{jp}^2 \left( \sum_{s=t-\ell}^{t-1} D_{js} \right)^2 \right)^{1/2}}{(np)^{1/2}} \right\|_{3a} \\
&\leq C \left( \ell^{1/2} \left( \frac{1}{n^{1/2}} + \frac{\overline{p}_n}{n} \right) + \frac{\overline{p}_n^{1/2} \ell}{n} + \frac{\ell^{1/2}}{n^{1/2}} \right) \leq C \left( \frac{\overline{p}_n}{n} \right)^{1/2}.
\end{aligned}$$

These bounds and (B.14) give, uniformly in $y$, $p$ and $t$,

$$\left\| \widetilde{\Sigma}_{pt}^{(1)}(y) \right\|_{3a} \leq \frac{C}{n^{1/2}}, \quad \left\| \mathbf{\Sigma}_{pt}^{(1)}(y) \right\|_{3a} \leq C \left( \frac{\overline{p}_n}{n} \right)^{1/2},$$
$$\left\| \widetilde{\Sigma}_{pt}^{(1,1)}(y) \right\|_{3a/2} \leq C \left( \frac{\overline{p}_n^{1/2}}{n} + \left( \frac{\overline{p}_n}{n} \right)^{3/2} + \frac{\overline{p}_n^{1/2}}{n^{3/2}} + \frac{1}{n \overline{p}_n^{1/2}} \right) \leq C \frac{\overline{p}_n^{1/2}}{n}.$$

Now, arguing as for the study of (B.29), $e = O\left(\overline{p}_n^{1/a}\right)$ give uniformly in $p$, $t$ and $y$,

$$\mathbb{E}\left[ \left| \mathfrak{I}_{1t}^{(1)}(y) \right| \right] + \mathbb{E}\left[ \left| \mathfrak{I}_{2t}^{(1)}(y) \right| \right] + \mathbb{E}\left[ \left| \mathfrak{I}_{4t}^{(1)}(y) \right| \right] \leq C \frac{\overline{p}_n^{1/2+3/a}}{n^{3/2}}, \quad \mathbb{E}\left[ \left| \mathfrak{I}_{3t}^{(1)}(y) \right| \right] \leq C \frac{\overline{p}_n^{1+3/a}}{n^{3/2}}.$$

It then follows $\sum_{t=1}^{n} \int_0^1 \left| \mathbb{E}\left[ \mathfrak{I}_t^{(1)}(y) \right] \right| dy \leq C \overline{p}_n^{1+3/a} / n^{1/2}$. Since $\sum_{t=1}^{n} \int_0^1 \left| \mathbb{E}\left[ \mathfrak{I}_t^{(1)}(y; \eta_t) \right] \right| dy$ satisfies a similar bound, we have for (B.34),

$$\sum_{t=1}^{n} \left| \int_0^1 \mathbb{E}\left[ \mathfrak{I}_t^{(1)}(y) - \mathfrak{I}_t^{(1)}(y; \eta_t) \right] dy \right| \leq C \frac{\overline{p}_n^{1+3/a}}{n^{1/2}}.$$

Consider now (B.33). Since $D_{jt}^{t-\ell+1}$ and $\eta_t$ are independent of the $\mathfrak{J}_{tn}^{(1)}(0)$, $\mathfrak{M}_t^{1-2e}(0)$ and $\mathbf{\Sigma}_{pt}(0)$, we have using (B.12),

$$
\begin{aligned}
&\mathbb{E}\left[\mathfrak{J}_t(0)-\mathfrak{J}_t(0;\eta_t)\right] \\
&= \frac{4}{n}\mathbb{E}\left[\left(\frac{1}{e}-1\right)\mathfrak{J}_{tn}^{(1)}(0)\,\mathfrak{M}_t^{1-2e}(0)\right. \\
&\sum_{p_1,p_2=1}^{\overline{p}} \mathbf{\Sigma}_{p_1t}^{e-1}(0)\,\mathbf{\Sigma}_{p_2t}^{e-1}(0)\, f\left(\mathbf{\Sigma}_{p_1t}^{e-1}(0)\right) f\left(\mathbf{\Sigma}_{p_2t}^{e-1}(0)\right)\sum_{j_1=1}^{p_1}\sum_{j_2=1}^{p_2}\left(\mathbb{E}\left[D_{j_1t}^{t-\ell+1}D_{j_2t}^{t-\ell+1}\right]-\mathbb{E}\left[\eta_{j_1t}\eta_{j_2t}\right]\right) \\
&\left.\frac{K_{j_1p_1}\left(\sum_{s_1=j_1+1}^{t-\ell+1}D_{j_1s_1}+\sum_{s_1=t-\ell}^{n}\eta_{j_1s_1}\right)}{n^{1/2}\sigma^4V_{\Delta}(p_1)}\frac{K_{j_2p_2}\left(\sum_{s_2=j_2+1}^{t-\ell+1}D_{j_2s_2}+\sum_{s_2=t-\ell}^{n}\eta_{j_2s_2}\right)}{n^{1/2}\sigma^4V_{\Delta}(p_2)}\right] \\
&= 0.
\end{aligned}
$$

Hence (B.33) and (B.34) give

$$
\left|\sum_{t=1}^{n}\mathbb{E}\left[\mathcal{I}_{1t}^{(2)}(0;D_t)-\mathcal{I}_{1t}^{(2)}(0;\eta_t)\right]\right|\leq C\frac{\overline{p}_n^{1+3/a}}{n^{1/2}}.
$$

To study $\left|\mathbb{E}\left[\mathcal{I}_{2t}^{(2)}(0;D_t)-\mathcal{I}_{2t}^{(2)}(0;\eta_t)\right]\right|$, observe that, uniformly with respect to $p$, $t$ and $y$,

$$
\max\left(\left\|\widetilde{\Sigma}_{pt}^{(2)}(y)\right\|_{3a/2},\left\|\widetilde{\Sigma}_{pt}^{(2)}(y;\eta_t)\right\|_{3a/2}\right)\leq C\frac{\overline{p}_n^{1/2}}{n},
$$

$$
\max\left(\left\|\widetilde{\Sigma}_{pt}^{(2,1)}(y)\right\|_{a},\left\|\widetilde{\Sigma}_{pt}^{(2,1)}(y;\eta_t)\right\|_{a}\right)\leq C\left(\frac{\overline{p}_n}{n^{3/2}}+\frac{1}{n\overline{p}_n}\right).
$$

Arguing as for $\sum_{t=1}^{n}\mathbb{E}\left[\mathcal{I}_{1t}^{(2)}(0;D_t)-\mathcal{I}_{1t}^{(2)}(0;\eta_t)\right]$ gives $\left|\sum_{t=1}^{n}\mathbb{E}\left[\mathcal{I}_{2t}^{(2)}(0;D_t)-\mathcal{I}_{2t}^{(2)}(0;\eta_t)\right]\right|\leq C\left(\frac{\overline{p}_n^{1+2/a}}{n^{1/2}}+\frac{\overline{p}_n^{1/a}}{\overline{p}_n}\right)$, and provided $e=O\left(\overline{p}_n^{1/(2a)}\right)$

$$
\left|\sum_{t=1}^{n}\mathbb{E}\left[\mathcal{I}_{3t}^{(2)}(0;D_t)-\mathcal{I}_{3t}^{(2)}(0;\eta_t)\right]\right|+\left|\sum_{t=1}^{n}\mathbb{E}\left[\mathcal{I}_{4t}^{(2)}(0;D_t)-\mathcal{I}_{4t}^{(2)}(0;\eta_t)\right]\right|\leq C\frac{\overline{p}_n^{1+3/a}}{n^{1/2}}.
$$

It then follows

$$
\left|\sum_{t=1}^{n}\mathbb{E}\left[\mathcal{I}_{t}^{(2)}(0;D_t)-\mathcal{I}_{t}^{(2)}(0;\eta_t)\right]\right|\leq C\left(\frac{\overline{p}_n^{1+3/a}}{n^{1/2}}+\frac{1}{\overline{p}_n^{1-1/a}}\right). \tag{B.35}
$$

Substituting (B.32), (B.35) in (B.29), (B.28) shows that the Lemma is proved. □

B.6.3. *End of the proof of Proposition A.3.* The rest of the proof is divided in 3 steps.

**Step 1: Martingale approximation**. Let $\widetilde{S}_p$ and $\check{S}_p$ be as in (A.1) and (B.16) respectively. Let $\mathfrak{a} = 4a/3$. The Cauchy-Schwarz inequality gives

$$\left|\check{S}_p - \widetilde{S}_p\right| = \sum_{j=1}^{p}\left(K_{jp}\frac{1}{n^{1/2}}\left|M_{jn} - \sum_{t=j+1}^{n} u_t u_{t-j}\right| \times \frac{1}{n^{1/2}}\left|M_{jn} + \sum_{t=j+1}^{n} u_t u_{t-j}\right|\right)$$
$$\leq C\left(\sum_{j=1}^{p}\frac{1}{n}\left(M_{jn} - \sum_{t=j+1}^{n} u_t u_{t-j}\right)^2\right)^{1/2}\left(\sum_{j=1}^{p}\frac{1}{n}\left(M_{jn} + \sum_{t=j+1}^{n} u_t u_{t-j}\right)^2\right)^{1/2}.$$

Hence

$$\left\|\check{S}_p - \widetilde{S}_p\right\|_{\mathfrak{a}/2}$$
$$\leq C\mathbb{E}^{\frac{1}{\mathfrak{a}}}\left[\left(\sum_{j=1}^{p}\frac{1}{n}\left(M_{jn} - \sum_{t=j+1}^{n} u_t u_{t-j}\right)^2\right)^{\frac{\mathfrak{a}}{2}}\right]\mathbb{E}^{\frac{1}{\mathfrak{a}}}\left[\left(\sum_{j=1}^{p}\frac{1}{n}\left(M_{jn} + \sum_{t=j+1}^{n} u_t u_{t-j}\right)^2\right)^{\frac{\mathfrak{a}}{2}}\right].$$

Observe now that (B.4) gives

$$\mathbb{E}^{\frac{1}{\mathfrak{a}}}\left[\left(\sum_{j=1}^{p}\frac{1}{n}\left(M_{jn} - \sum_{t=j+1}^{n} u_t u_{t-j}\right)^2\right)^{\frac{\mathfrak{a}}{2}}\right]$$
$$\leq \left(\frac{1}{n}\sum_{j=1}^{p}\mathbb{E}^{\frac{2}{\mathfrak{a}}}\left[\left|M_{jn} - \sum_{t=j+1}^{n} u_t u_{t-j}\right|^{\mathfrak{a}}\right]\right)^{1/2} \leq C\left(\frac{p}{n}\right)^{1/2}.$$

Since the Burkholder inequality and $\max_j \mathbb{E}\left[|D_{jt}|^{\mathfrak{a}}\right] < \infty$ give $\max_{j\in[1,\overline{p}_n]} \mathbb{E}^{1/\mathfrak{a}}\left[|M_{jn}|^{\mathfrak{a}}\right] \leq Cn^{1/2}$, we also have

$$\begin{aligned}
&\mathbb{E}^{\frac{1}{\mathfrak{a}}}\left[\left(\left(\sum_{j=1}^{p}\frac{1}{n}\left(M_{jn}+\sum_{t=j+1}^{n}u_t u_{t-j}\right)^2\right)\right)^{\frac{\mathfrak{a}}{2}}\right] \\
&\leq \left(\frac{1}{n}\sum_{j=1}^{p}\left(\mathbb{E}^{\frac{1}{\mathfrak{a}}}\left[\left|M_{jn}+\sum_{t=j+1}^{n}u_t u_{t-j}\right|^{\mathfrak{a}}\right]\right)^2\right)^{1/2} \\
&\leq \left(\frac{1}{n}\sum_{j=1}^{p}\left(2\mathbb{E}^{\frac{1}{\mathfrak{a}}}\left[|M_{jn}|^{\mathfrak{a}}\right]+\mathbb{E}^{\frac{1}{\mathfrak{a}}}\left[\left|\sum_{t=j+1}^{n}u_t u_{t-j}-M_{jn}\right|^{\mathfrak{a}}\right]\right)^2\right)^{1/2} \\
&\leq \left(\frac{p(Cn^{1/2}+C)^2}{n}\right)^{1/2} \leq Cp^{1/2}.
\end{aligned}$$

It then follows that $\left\|\check{S}_p - \widetilde{S}_p\right\|_{\mathfrak{a}/2} \leq Cp/n^{1/2}$ and them $\max_{p\in[1,\overline{p}_n]} \mathbb{E}\left[\left|\left(\check{S}_p - \widetilde{S}_p\right)/p^{1/2}\right|^{\mathfrak{a}/2}\right] \leq C\left(\overline{p}_n/n\right)^{\mathfrak{a}/4}$. Hence the Markov inequality gives

$$\begin{aligned}
\mathbb{P}\left(\max_{p\in[1,\overline{p}_n]}\left|\frac{\check{S}_p-\widetilde{S}_p}{p^{1/2}}\right| \geq t\right) &\leq \sum_{p=1}^{\overline{p}_n}\mathbb{P}\left(\left|\frac{\check{S}_p-\widetilde{S}_p}{p^{1/2}}\right| \geq t\right) \\
&\leq \frac{\overline{p}_n}{t^{\mathfrak{a}/2}}\max_{p\in[1,\overline{p}_n]}\mathbb{E}\left[\left|\frac{\check{S}_p-\widetilde{S}_p}{p^{1/2}}\right|^{\frac{\mathfrak{a}}{2}}\right] \leq \frac{C}{t^{a/2}}\left(\frac{\overline{p}_n^{1+\frac{4}{\mathfrak{a}}}}{n}\right)^{\mathfrak{a}/4},
\end{aligned}$$

and $\overline{p}_n = o\left(n^{1/(2(1+4/\mathfrak{a}))}\right)$ gives

$$\max_{p\in[1,\overline{p}_n]}\left|\frac{\check{S}_p-\widetilde{S}_p}{p^{1/2}}\right| = o_{\mathbb{P}}(1). \tag{B.36}$$

**Step 2: some Gaussian approximations**. Let $\gamma_n' = \gamma_n\left(1+\epsilon/2\right)/\left(1+\epsilon\right)$. (3.1) gives $\gamma_n \geq \gamma_n' \geq \widetilde{\gamma}_n = \left(2\ln\ln\overline{p}_n\right)^{1/2}\left(1+\epsilon/3\right)$. Consider a three times continuously differentiable function $\iota\left(x\right)$ with $\max_{j=1,2,3}\sup_x\left|\iota^{(3)}\left(x\right)\right| < \infty$ and $\mathbb{I}\left(x\geq 0\right) \leq \iota\left(x\right) \leq \mathbb{I}\left(x \geq -\epsilon\right)$. Let $\mathcal{I}\left(x\right) = \iota\left(x-\gamma_n'\right)$. Let $\check{s}_p$ be as in (B.16). Then Lemma B.5 with $e = \overline{p}_n^{1/(2a)}$, (B.14) and

(B.16), and Assumption R give

$$\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\{\check{s}_p\}\geq\gamma_n'\right)\leq\mathbb{P}\left(\mathcal{M}\geq\gamma_n'\right)\leq\mathbb{E}\left[\mathcal{I}\left(\mathcal{M}\right)\right]$$
$$\leq\mathbb{E}\left[\mathcal{I}\left(\mathcal{M}_1\left(\eta_1\right)\right)\right]+o\left(1\right)\leq\mathbb{P}\left(\mathcal{M}_1\left(\eta_1\right)\geq\gamma_n'-\epsilon\right)+o\left(1\right).$$

We now look for a more explicit expression for the RHS. Recall that $\mathcal{M}_1\left(\eta_1\right)=\left(\sum_{p=1}^{\overline{p}_n}f^e\left(\check{s}_{p1}\left(1;\eta_1\right)\right)\right)^{1/e}$. Consider $\Omega\left(p\right)=\left[\omega_1,\ldots,\omega_p\right]'$ where the $\omega_p$'s are i.i.d. standard normal variables,

$$\mathcal{K}\left(p\right)=\operatorname{Diag}\left(\left(1-j/n\right)K_{jp},j=1,\ldots,p\right),$$
$$\mathcal{C}_\eta\left(p\right)=\left[\operatorname{Cov}\left(\eta_{j_1t},\eta_{j_2t}\right),j_1,j_2=1,\ldots,p\right],$$
$$\mathcal{V}_\eta\left(p\right)=\mathcal{C}_\eta^{1/2}\left(p\right)\mathcal{K}\left(p\right)\mathcal{C}_\eta^{1/2}\left(p\right),$$

and $\mathcal{D}_\eta\left(p\right)=\operatorname{Diag}\left(\left(1-j/n\right)K_{jp}\operatorname{Var}\left(\eta_{jt}\right),j=1,\ldots,p\right)$ the $p\times p$ diagonal matrix obtained from the diagonal entries of $\mathcal{V}_\eta\left(p\right)$. Then the $\check{s}_{p1}\left(1;\eta_1\right)$, $p=1,\ldots,\overline{p}_n$, have the same joint distribution than

$$\tilde{s}_p=\frac{\Omega\left(p\right)'\mathcal{V}_\eta\left(p\right)\Omega\left(p\right)-\sigma^4E_\Delta\left(p\right)}{\sigma^4V_\Delta\left(p\right)},\quad p=1,\ldots,\overline{p}_n,$$

so that $\mathcal{M}_1\left(\eta_1\right)$ and $\widetilde{\mathcal{M}}=\left(\sum_{p=1}^{\overline{p}_n}f^e\left(\tilde{s}_p\right)\right)^{1/e}$ have the same distribution, and then

$$\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\{\check{s}_p\}\geq\gamma_n'\right)\leq\mathbb{P}\left(\widetilde{\mathcal{M}}\geq\gamma_n'-\epsilon\right)+o\left(1\right).$$

Define now

$$\bar{s}_p=\frac{\Omega\left(p\right)'\mathcal{D}_\eta\left(p\right)\Omega\left(p\right)-\sigma^4E_\Delta\left(p\right)}{\sigma^4V_\Delta\left(p\right)}=\frac{\sum_{j=1}^{p}\left(1-\frac{j}{n}\right)K_{jp}\operatorname{Var}\left(\eta_{jt}\right)\omega_j^2-\sigma^4E_\Delta\left(p\right)}{\sigma^4V_\Delta\left(p\right)}.$$

Then for all $p = 1, \ldots, \overline{p}_n$,

$$\begin{aligned}
|\tilde{s}_p - \bar{s}_p| &= \left| \frac{\Omega(p)' (\mathcal{V}_\eta(p) - \mathcal{D}_\eta(p)) \Omega(p)}{\sigma^4 V_\Delta(p)} \right| \\
&\leq C \sum_{1 \leq j_1 \neq j_2 \leq p} \left| \operatorname{Cov}\left( \left(1 - \frac{j_1}{n}\right)^{1/2} K_{j_1 p}^{1/2} \eta_{j_1 t}, \left(1 - \frac{j_2}{n}\right)^{1/2} K_{j_2 p}^{1/2} \eta_{j_2 t} \right) \right| |\omega_{j_1}| \, |\omega_{j_2}| \\
&\leq C \sum_{1 \leq j_1 \neq j_2 \leq \overline{p}_n} |\operatorname{Cov}(\eta_{j_1 t}, \eta_{j_2 t})| \, |\omega_{j_1}| \, |\omega_{j_2}| = O_{\mathbb{P}}(1),
\end{aligned}$$

by Lemma B.3. Hence since $f(x) \leq 2 \vee x$ by (B.14) and using (B.15),

$$\begin{aligned}
\widetilde{\mathcal{M}} &\leq \left(1 + O\left(\frac{\ln n}{\overline{p}_n^{1/(2a)}}\right)\right) \max_{p \in [2, \overline{p}_n]} \{2 \vee \tilde{s}_p\} \leq \left(1 + O\left(\frac{\ln n}{\overline{p}_n^{1/(2a)}}\right)\right) 2 \vee \max_{p \in [2, \overline{p}_n]} \{\tilde{s}_p\} \\
&\leq \left(1 + O\left(\frac{\ln n}{n^{1/8a}}\right)\right) \max_{p \in [2, \overline{p}_n]} \{\bar{s}_p\} + O_{\mathbb{P}}(1).
\end{aligned}$$

Define now

$$\mathsf{V}_\Delta(p) = \left(2 \sum_{j=1}^{p} K_{jp}^2\right)^{1/2}, \quad \mathsf{s}_p = \frac{\sum_{j=1}^{p} K_{jp} \left(\omega_j^2 - 1\right)}{\mathsf{V}_\Delta(p)},$$

which is such that

$$\begin{aligned}
&|\bar{s}_p - \mathsf{s}_p| \leq |\mathfrak{e}_{1p}| + |\mathfrak{e}_{2p}| \text{ where} \\
&\mathfrak{e}_{1p} = \left(\frac{\sigma^4 \mathsf{V}_\Delta(p)}{\sigma^4 V_\Delta(p)} - 1\right) \sigma^4 \mathsf{s}_p, \\
&\mathfrak{e}_{2p} = \frac{\sum_{j=1}^{p} \left\{\left(1 - \frac{j}{n}\right) \operatorname{Var}(\eta_{jt}) - \sigma^4\right\} K_{jp} \omega_j^2 - \sigma^4 \sum_{j=1}^{p} \frac{j}{n} K_{jp}}{\sigma^4 V_\Delta(p)}.
\end{aligned}$$

Since $K'(\cdot)$ is continuous on $[0,1]$, the Weierstrass Theorem implies it can be uniformly approximated with a sequence of polynomial function. Hence (B.1), Assumption K and the LIL for weighted sums in Li and Tomkins (1996) gives that

$$\limsup_{p \to \infty} \frac{|\mathsf{V}_\Delta(p) \, \mathsf{s}_p|}{p^{1/2} (2 \ln \ln p)^{1/2}} \leq \left(2 \int K^4(t) \, dt\right)^{1/2}, \text{ almost surely.}$$

Since, under Assumption K, $\mathsf{V}_{\Delta}(p)/p^{1/2} \to \left(2\int K^4(t)\,dt\right)^{1/2}$ by convergence of Riemann sums, this gives

$$\sup_{p\in[2,\overline{p}_n]} |\mathsf{s}_p| \le (2\ln\ln\overline{p}_n)^{1/2}(1+o_{\mathbb{P}}(1)). \tag{B.37}$$

Observe also that Lemma A.2-(ii), $\overline{p}_n = o\left(n^{1/2}\right)$, and Assumption K give uniformly in $p\in[1,\overline{p}_n]$

$$\left|\frac{\mathsf{V}_{\Delta}(p)}{V_{\Delta}(p)} - 1\right| \le C\left(\frac{1}{p}\sum_{j=1}^{p}\frac{j^2}{n^2}K_{jp}^2\right)^{1/2} = o\left(\frac{1}{n^{1/2}}\right).$$

Hence

$$\max_{p\in[2,\overline{p}_n]}|\mathfrak{e}_{1p}| = o_{\mathbb{P}}\left(\left(\frac{2\ln\ln\overline{p}_n}{n}\right)^{1/2}\right) = o_{\mathbb{P}}(1).$$

Now, for $\max_{p\in[2,\overline{p}_n]}|\mathfrak{e}_{2p}|$, we have by Lemmas A.2-(ii) and B.3, $\overline{p}_n = o\left(n^{1/2}\right)$, and Assumption K,

$$\max_{p\in[2,\overline{p}_n]}|\mathfrak{e}_{2p}| \le C\left\{\sum_{j=1}^{\overline{p}_n}\left|\mathrm{Var}(\eta_{jt})-\sigma^4\right|\omega_j^2 + \frac{1}{n}\sum_{j=1}^{\overline{p}_n} j\omega_j^2 + \frac{\overline{p}_n^{3/2}}{n}\right\} = O_{\mathbb{P}}(1) + O_{\mathbb{P}}\left(\frac{\overline{p}_n^2}{n}\right) = O_{\mathbb{P}}(1).$$

Hence $\max_{p\in[2,\overline{p}_n]}|\bar{s}_p - \mathsf{s}_p| = O_{\mathbb{P}}(1)$ and substituting in the bounds for $\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\{\check{s}_p\}\ge\gamma_n'\right)$ and $\widetilde{\mathcal{M}}$ above gives, by (3.1), $\gamma_n' = \gamma_n(1+\epsilon/2)/(1+\epsilon)$, $\gamma_n' \ge (2\ln\ln\overline{p}_n)^{1/2}(1+\epsilon/3)$ and (B.37)

$$\begin{aligned}
\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\{\check{s}_p\}\ge\gamma_n'\right) &= \mathbb{P}\left(\left(1+O\left(\frac{\ln n}{n^{1/8a}}\right)\right)\max_{p\in[2,\overline{p}_n]}\{\mathsf{s}_p\} + O_{\mathbb{P}}(1) \ge \gamma_n' - \epsilon\right) + o(1) \\
&\le \mathbb{P}\left(\max_{p\in[2,\overline{p}_n]}\{\mathsf{s}_p\} \ge (2\ln\ln\overline{p}_n)^{1/2}(1+\epsilon/3)\right) + o(1) \\
&= o(1).
\end{aligned} \tag{B.38}$$

**Step 3: Conclusion**. Propositions A.2 and A.1, Lemma A.2 and $\overline{p}_n = O\left(n^{1/2}\right)$, the expression of $\check{S}_p$ and $\check{s}_p$ in (B.16) and (B.36) gives

$$\begin{aligned}
&\max_{p\in[2,\overline{p}_n]} \frac{\left(\widehat{S}_p - \widehat{S}_1\right)/\widehat{R}_0^2 - E_\Delta(p)}{V_\Delta(p)} = \max_{p\in[2,\overline{p}_n]} \frac{\left(\widehat{S}_p - \widehat{S}_1\right) - \widehat{R}_0^2 E_\Delta(p)}{\widehat{R}_0^2 V_\Delta(p)} \\
&= (1 + o_{\mathbb{P}}(1)) \max_{p\in[2,\overline{p}_n]} \frac{\left(\widetilde{S}_p - \widetilde{S}_1\right) - R_0^2 E_\Delta(p)}{R_0^2 V_\Delta(p)} + O_{\mathbb{P}}\left(1 + \overline{p}_n^{1/2}\left(\widehat{R}_0^2 - R_0^2\right)\right) \\
&= (1 + o_{\mathbb{P}}(1)) \max_{p\in[2,\overline{p}_n]} \{\check{s}_p\} + O_{\mathbb{P}}(1).
\end{aligned}$$

Hence (B.38) gives, since $\gamma_n - \gamma_n' \to +\infty$,

$$\mathbb{P}\left(\max_{p\in[2,\overline{p}_n]} \frac{\left(\widehat{S}_p - \widehat{S}_1\right)/\widehat{R}_0^2 - E_\Delta(p)}{V_\Delta(p)} \geq \gamma_n\right) \leq \mathbb{P}\left(\max_{p\in[2,\overline{p}_n]} \{\check{s}_p\} \geq \gamma_n'\right) + o(1) = o(1).$$

This ends the proof of the Proposition. □

B.7. **Proof of Propositions A.4 and A.5.** When studying the mean and variance of $\widetilde{S}_p$, we make use of Theorem 2.3.2 in Brillinger (2001) which implies in particular that, for any real zero-mean random variables $Z_1, \ldots, Z_4$,

$$\begin{aligned}
&\operatorname{Var}(Z_1 Z_2, Z_3 Z_4) = \operatorname{Var}(Z_1, Z_3)\operatorname{Var}(Z_2, Z_4) + \operatorname{Var}(Z_1, Z_4)\operatorname{Var}(Z_2, Z_3) \\
&+ \operatorname{Cum}(Z_1, Z_2, Z_3, Z_4). \qquad \text{(B.39)}
\end{aligned}$$

Note that Assumption R and Theorem B.1 imply that

$$\sup_{n,q\in[2,8]} \sum_{t_2,\ldots,t_q=-\infty}^{\infty} |\Gamma_n(0, t_2, \ldots, t_q)| < \infty. \qquad \text{(B.40)}$$

B.7.1. ***Proof of Proposition A.4***. (B.39) yields

$$\begin{aligned}
&\mathbb{E}\left[\widetilde{R}_j^2\right] = \frac{1}{n^2} \sum_{t_1,t_2=1}^{n-j} \mathbb{E}\left[u_{t_1} u_{t_1+j} u_{t_2} u_{t_2+j}\right] \\
&= \frac{1}{n^2} \sum_{t_1,t_2=1}^{n-j} \left(R_j^2 + R_{t_2-t_1}^2 + R_{t_2-t_1+j} R_{t_2-t_1-j} + \Gamma(0, j, t_2 - t_1, t_2 - t_1 + j)\right),
\end{aligned}$$

where

$$\sum_{t_1,t_2=1}^{n-j} R_{t_2-t_1}^2 = (n-j)R_0^2 + 2\sum_{\ell=1}^{n-j-1}(n-j-\ell)R_\ell^2,$$

$$\sum_{t_1,t_2=1}^{n-j} R_{t_2-t_1+j}R_{t_2-t_1-j} = (n-j)R_j^2 + 2\sum_{\ell=1}^{n-j-1}(n-j-\ell)R_{\ell+j}R_{\ell-j},$$

$$\sum_{t_1,t_2=1}^{n-j} \Gamma\left(0,j,t_2-t_1,t_2-t_1+j\right) = \sum_{\ell=-n+j+1}^{n-j-1}(n-j-|\ell|)\,\Gamma\left(0,j,\ell,\ell+j\right).$$

Set $k_j = K^2\left(j/p\right)$ to prove the first equality and $k_j = K^2\left(j/p\right)/\tau_j^2$ for the second. Note that Assumptions K and R give, in both case, $\max_{j\in[1,n-1]} k_j \leq C$ and $k_j \geq C\mathbb{I}\left(j \leq p/2\right)$. The equalities above give

$$\begin{aligned}
&\mathbb{E}\left[\sum_{j=1}^{n-1} k_j \widetilde{R}_j^2\right] - R_0^2\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)k_j \\
&= n\sum_{j=1}^{n-1}\left(\left(1-\frac{j}{n}\right)^2 + \frac{1}{n}\left(1-\frac{j}{n}\right)\right)k_j R_j^2 \qquad \text{(B.41)}\\
&+ 2\sum_{j=1}^{n-1} k_j \sum_{\ell=1}^{n-j-1}\left(1-\frac{j+\ell}{n}\right)\left(R_\ell^2 + R_{\ell+j}R_{\ell-j}\right)\\
&+ \sum_{j=1}^{n-1} k_j \sum_{\ell=-n+j+1}^{n-j-1}\left(1-\frac{j+|\ell|}{n}\right)\Gamma\left(0,j,\ell,\ell+j\right).
\end{aligned}$$

We start with the item $R_0^2\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)k_j$, which is equal to $R_0^2 E\left(p\right)$ when $k_j = K^2\left(j/p\right)$, that is when proving the first equality. When $k_j = K^2\left(j/p\right)/\tau_j^2$, (A.4) gives, under Assumptions K and R,

$$\left|R_0^2\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)k_j - E\left(p\right)\right| \leq C\sum_{j=1}^{p}\left|\tau_j^2 - R_0^2\right| \leq C\sum_{j=1}^{\infty} j^{-6}$$

so that $R_0^2\sum_{j=1}^{n-1}\left(1-j/n\right)k_j \geq E\left(p\right) - C'$.

Let us now turn to the other items. The lower bound$k_j \geq CI(j \leq p/2)$ gives that (B.41) is larger than $Cn\sum_{j=1}^{p/2} R_j^2$. To bound the remaining terms in (B.41), we note that by

Assumptions K, R and (B.40),

$$\left|\sum_{j=1}^{n-1} k_j \sum_{\ell=1}^{n-j-1}\left(1-\frac{j+\ell}{n}\right) R_\ell^2\right| \leq C \sum_{j=1}^{n-1} \mathbb{I}(j \leq p) \times \sum_{j=1}^{\infty} R_j^2 \leq C p \sum_{j=1}^{\infty} R_j^2 = o(n) \sum_{j=1}^{\infty} R_j^2,$$

$$\left|\sum_{j=1}^{n-1} k_j \sum_{\ell=1}^{n-j-1}\left(1-\frac{j+\ell}{n}\right) R_{\ell+j} R_{\ell-j}\right| \leq C \sum_{j=1}^{+\infty} \sum_{\ell=1}^{+\infty} \left|R_{\ell+j} R_{\ell-j}\right| \leq C\left(\sum_{j=0}^{\infty}\left|R_j\right|\right)^2 \leq C,$$

$$\left|\sum_{j=1}^{n-1} k_j \sum_{\ell=-n+j+1}^{n-j-1}\left(1-\frac{j+\ell}{n}\right) \Gamma\left(0, j, \ell, \ell+j\right)\right| \leq C \sum_{t_2, t_3, t_4=-\infty}^{\infty} \left|\Gamma(0, t_2, t_3, t_4)\right| \leq C$$

uniformly with respect to $p \in [1, \overline{p}_n]$. Substituting these bounds in the equality above establishes the proposition. □

B.7.2. ***Proof of Proposition A.5***. Let $f$ be the spectral density of the alternative. Using (B.40), we obtain

$$\sup_{\lambda \in [-\pi, \pi]} |f(\lambda)| \leq C \quad \text{and} \quad \sum_{j=1}^{\infty} R_j^2 \leq C \tag{B.42}$$

because $\sup_{\lambda \in [-\pi, \pi]} |f(\lambda)| \leq \left(|R_0| + 2\sum_{j=1}^{\infty} |R_j|\right)/(2\pi)$ and $\sum_{j=1}^{\infty} R_j^2 \leq \left(\sum_{j=1}^{\infty} |R_j|\right)^2$. We recall that $\widetilde{R}_j = \sum_{t=1}^{n-j} u_t u_{t+j}/n$ and define $\overline{R}_j = \mathbb{E}\left[\widetilde{R}_j\right] = (1 - j/n) R_j$. Set $k_j = K^2(j/p)$ to prove the first equality and $k_j = K^2(j/p)/\tau_j^2$ for the second. Note that Assumptions K and R give, in both case, $k_j \leq C\mathbb{I}(j \leq p)$. To avoid notation burdens, redefine $\widetilde{S}_p$ as $\sum_{j=1}^{n-1} k_j \widetilde{R}_j^2$. Define $D_j = \widetilde{R}_j - \overline{R}_j$. We have $\mathbb{E}[D_j] = 0$ and $\widetilde{S}_p = n\sum_{j=1}^{n-1} k_j \overline{R}_j^2 + 2n\sum_{j=1}^{n-1} k_j \overline{R}_j D_j + n\sum_{j=1}^{n-1} k_j D_j^2$. The inequality $(a+b)^2 \leq 2a^2 + 2b^2$ implies that

$$\operatorname{Var}\left(\widetilde{S}_p\right) \leq 4 \operatorname{Var}\left(n \sum_{j=1}^{n-1} k_j \overline{R}_j \widetilde{R}_j\right) + 2 \operatorname{Var}\left(n \sum_{j=1}^{n-1} k_j D_j^2\right). \tag{B.43}$$

By identity (B.39),

$$\operatorname{Var}\left(n \sum_{j=1}^{n-1} k_j \overline{R}_j \widetilde{R}_j\right) = \sum_{j_1, j_2=1}^{n-1} k_{j_1} k_{j_2} \overline{R}_{j_1} \overline{R}_{j_2} \sum_{t_1=1}^{n-j_1} \sum_{t_2=1}^{n-j_2} \operatorname{Cov}\left(u_{t_1} u_{t_1+j_1}, u_{t_2} u_{t_2+j_2}\right) \leq V_1 + K_1$$

with

$$V_1 = \left| \sum_{j_1, j_2 = 1}^{n-1} k_{j_1} k_{j_2} \overline{R}_{j_1} \overline{R}_{j_2} \sum_{t_1=1}^{n-j_1} \sum_{t_2=1}^{n-j_2} \left( R_{t_2 - t_1} R_{t_2 - t_1 + j_2 - j_1} + R_{t_2 - t_1 - j_1} R_{t_2 - t_1 + j_2} \right) \right|,$$

$$K_1 = \left| \sum_{j_1, j_2 = 1}^{n-1} k_{j_1} k_{j_2} \overline{R}_{j_1} \overline{R}_{j_2} \sum_{t_1=1}^{n-j_1} \sum_{t_2=1}^{n-j_2} \Gamma \left( t_1, t_1 + j_1, t_2, t_2 + j_2 \right) \right|.$$

The second term on the right of (B.43) is, up to a multiplicative constant, equal to

$$\operatorname{Var}\left( n \sum_{j=1}^{n-1} k_j D_j^2 \right) = n^2 \sum_{j_1, j_2 = 1}^{n-1} k_{j_1} k_{j_2} \operatorname{Cov}\left( D_{j_1}^2, D_{j_2}^2 \right).$$

Applying (B.39) twice we obtain

$$\begin{aligned}
&\operatorname{Cov}\left( D_{j_1}^2, D_{j_2}^2 \right) \\
&= \frac{1}{n^4} \sum_{t_1, t_2 = 1}^{n-j_1} \sum_{t_3, t_4 = 1}^{n-j_2} \operatorname{Cov}\left[ \prod_{q=1}^{2} \left( u_{t_q} u_{t_q + j_1} - \mathbb{E}[u_{t_q} u_{t_q + j_1}] \right), \prod_{q=3}^{4} \left( u_{t_q} u_{t_q + j_2} - \mathbb{E}[u_{t_q} u_{t_q + j_2}] \right) \right] \\
&= \frac{1}{n^4} \sum_{t_1, t_2 = 1}^{n-j_1} \sum_{t_3, t_4 = 1}^{n-j_2} [\operatorname{Cov}\left( u_{t_1} u_{t_1 + j_1}, u_{t_3} u_{t_3 + j_2} \right) \operatorname{Cov}\left( u_{t_2} u_{t_2 + j_1}, u_{t_4} u_{t_4 + j_2} \right) \\
&\qquad + \operatorname{Cov}\left( u_{t_1} u_{t_1 + j_1}, u_{t_4} u_{t_4 + j_2} \right) \operatorname{Cov}\left( u_{t_2} u_{t_2 + j_1}, u_{t_3} u_{t_3 + j_2} \right)] \\
&+ \frac{1}{n^4} \sum_{t_1, t_2 = 1}^{n-j_1} \sum_{t_3, t_4 = 1}^{n-j_2} \operatorname{Cum}\left( u_{t_1} u_{t_1 + j_1}, u_{t_2} u_{t_2 + j_1}, u_{t_3} u_{t_3 + j_2}, u_{t_4} u_{t_4 + j_2} \right) \\
&= \frac{2}{n^4} \left( \sum_{t_1=1}^{n-j_1} \sum_{t_2=1}^{n-j_2} \left( R_{t_2 - t_1} R_{t_2 - t_1 + j_2 - j_1} + R_{t_2 - t_1 - j_1} R_{t_2 - t_1 + j_2} + \Gamma(t_1, t_1 + j_1, t_2, t_2 + j_2) \right) \right)^2 \\
&+ \frac{1}{n^4} \sum_{t_1, t_2 = 1}^{n-j_1} \sum_{t_3, t_4 = 1}^{n-j_2} \operatorname{Cum}\left( u_{t_1} u_{t_1 + j_1}, u_{t_2} u_{t_2 + j_1}, u_{t_3} u_{t_3 + j_2}, u_{t_4} u_{t_4 + j_2} \right).
\end{aligned}$$

Since $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$, we can write $\mathrm{Var}\left(n\sum_{j=1}^{n-1} k_j D_j^2\right) \leq 6V_2 + K_2 + 6K_2'$ with

$$V_2 = \frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\left(\left(\sum_{t_1=1}^{n-j_1}\sum_{t_2=1}^{n-j_2} R_{t_2-t_1}R_{t_2-t_1+j_2-j_1}\right)^2 + \left(\sum_{t_1=1}^{n-j_1}\sum_{t_2=1}^{n-j_2} R_{t_2-t_1-j_1}R_{t_2-t_1+j_2}\right)^2\right),$$

$$K_2 = \left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2} \mathrm{Cum}\left(u_{t_1}u_{t_1+j_1}, u_{t_2}u_{t_2+j_1}, u_{t_3}u_{t_3+j_2}, u_{t_4}u_{t_4+j_2}\right)\right|,$$

$$K_2' = \frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\left(\sum_{t_1=1}^{n-j_1}\sum_{t_2=1}^{n-j_2}\Gamma\left(t_1, t_1+j_1, t_2, t_2+j_2\right)\right)^2,$$

Substituting in (B.43) shows that the proposition holds if the following inequalities hold:

$$V_1 \leq Cn\sum_{j=1}^{p} R_j^2, \quad V_2 \leq Cp, \quad K_1 \leq C, \quad K_2' \leq C, \quad K_2 \leq C\frac{p^2}{n}.$$

We establish these inequalities in five steps.

*Step 1: bound for $V_1$.* We note that $|\overline{R}_j| \leq |R_j|$ and that under Assumption K, $0 \leq k_j \leq C$ for all $j$. Using a covariance spectral representation $R_j = \int_{-\pi}^{\pi}\exp(\pm ij\lambda)f(\lambda)d\lambda$, the Cauchy-Schwarz inequality and (B.42), we obtain by Assumption K

$$\begin{aligned}
&\left|\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\overline{R}_{j_1}\overline{R}_{j_2}\sum_{t_1=1}^{n-j_1}\sum_{t_2=1}^{n-j_2} R_{t_2-t_1}R_{t_2-t_1+j_2-j_1}\right| \\
&= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left|\sum_{j=1}^{n-1} k_j\overline{R}_j\sum_{t=1}^{n-j} \mathrm{e}^{it\lambda_1}\mathrm{e}^{i(t+j)\lambda_2}\right|^2 f(\lambda_1)f(\lambda_2)d\lambda_1 d\lambda_2 \\
&\leq \left(\sup_{\lambda\in[-\pi,\pi]}|f(\lambda)|\right)^2\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j_1,j_2=1}^{n-1} k_{j_1}\overline{R}_{j_1}k_{j_2}\overline{R}_{j_2}\sum_{t_1=1}^{n-j_1}\sum_{t_2=1}^{n-j_2}\mathrm{e}^{it_1\lambda_1}\mathrm{e}^{i(t_1+j_1)\lambda_2}\mathrm{e}^{-it_2\lambda_1}\mathrm{e}^{-i(t_2+j_2)\lambda_2}d\lambda_1 d\lambda_2 \\
&\leq C\sum_{j=1}^{n-1}(n-j)k_j^2\overline{R}_j^2 \leq Cn\sum_{j=1}^{p} R_j^2,
\end{aligned}$$

$$\left|\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\overline{R}_{j_1}\overline{R}_{j_2}\sum_{t_1=1}^{n-j_1}\sum_{t_2=1}^{n-j_2} R_{t_2-t_1-j_1}R_{t_2-t_1+j_2}\right|$$

$$= \left|\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j_1=1}^{n-1} k_{j_1}\overline{R}_{j_1}\sum_{t_1=1}^{n-j_1} \mathrm{e}^{-i(t_1+j_1)\lambda_1}\mathrm{e}^{-it_1\lambda_2} \times \sum_{j_2=1}^{n-1} k_{j_2}\overline{R}_{j_2}\sum_{t_2=1}^{n-j_2} \mathrm{e}^{it_2\lambda_1}\mathrm{e}^{i(t_2+j_2)} f(\lambda_1)f(\lambda_2)d\lambda_1 d\lambda_2\right|$$

$$\leq \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left|\sum_{j=1}^{n-1} k_j\overline{R}_j\sum_{t=1}^{n-j} \mathrm{e}^{it\lambda_1}\mathrm{e}^{i(t+j)\lambda_2}\right|^2 f(\lambda_1)f(\lambda_2)d\lambda_1 d\lambda_2 \leq Cn\sum_{j=1}^{p} R_j^2$$

This establishes the bound for $V_1$.

*Step 2: bound for* $V_2$. We define $t_2 = t_1 + t_2'$, $j_2 = j_1 + j_2'$. By Assumption K and by (B.40),

$$\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\left(\sum_{t_1=1}^{n-j_1}\sum_{t_2=1}^{n-j_2} R_{t_2-t_1}R_{t_2-t_1-j_1+j_2}\right)^2$$

$$\leq \frac{C}{n^2}\sum_{j_1=1}^{n-1} K^2(j_1/p)\sum_{j_2\prime=-\infty}^{\infty}\left(n\sum_{t_2\prime=-\infty}^{+\infty}|R_{t_2\prime}R_{t_2\prime+j_2\prime}|\right)^2$$

$$\leq Cp \times \left(\sum_{j_2,t_1,t_2=-\infty}^{\infty}|R_{t_1}R_{t_1+j_2}R_{t_2}R_{t_2+j_2}|\right) \leq Cp\left(\sum_{t=-\infty}^{\infty}|R_t|\right)^4 \leq Cp,$$

$$\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\left(\sum_{t_1=1}^{n-j_1}\sum_{t_2=1}^{n-j_2} R_{t_2-t_1-j_1}R_{t_2-t_1+j_2}\right)^2$$

$$\leq \frac{C}{n^2}\sum_{j_1=1}^{n-1} K^2(j_1/p)\sum_{j_2\prime=-\infty}^{\infty}\left(n\sum_{t_2\prime=-\infty}^{+\infty}|R_{t_2\prime-j_1}R_{t_2\prime+j_1+j_2\prime}|\right)^2$$

$$\leq Cp\sum_{j_2',t_1,t_2=-\infty}^{\infty}\left|R_{t_1-j_1}R_{t_1+j_1+j_2'}R_{t_2-j_1}R_{t_2+j_1+j_2'}\right| \leq Cp\sum_{j,t_1,t_2=-\infty}^{\infty}|R_{t_1}R_{t_1+j}R_{t_2}R_{t_2+j}|$$

$$\leq Cp\left(\sum_{t=-\infty}^{\infty}|R_t|\right)^4 \leq Cp,$$

therefore $V_2 \leq Cp$.

*Step 3: bound for* $K_1$. Define $t_2 = t_1 + t$. Assumption K, and (B.40) yield

$$K_1 \leq C\sum_{j_1,j_2=1}^{p}\sum_{t=-\infty}^{\infty}|\Gamma(0,j_1,t,t+j_2)| \leq \sum_{t_1,t_2,t_3=-\infty}^{\infty}|\Gamma(0,t_1,t_2,t_3)|\,.$$

*Step 4: bound for $K_2'$.* (B.40) gives

$$
\begin{aligned}
K_2' &\leq \frac{1}{n^2} \sum_{j_1,j_2=1}^{n-1} k_{j_1} k_{j_2} \left( \sum_{t_1=1}^{n-j_1} \sum_{t_2=1}^{n-j_2} \left| \Gamma\left(0, j_1, t_2 - t_1, t_2 - t_1 + j_2\right) \right| \right)^2 \\
&\leq C \sum_{j_1,j_2=1}^{+\infty} \left( \sum_{t=-\infty}^{\infty} \left| \Gamma(0, j_1, t, t + j_2) \right| \right)^2 \\
&= C \sum_{j_1,j_2=1}^{+\infty} \sum_{t_1,t_2=-\infty}^{\infty} \left| \Gamma(0, j_1, t_1, t_1 + j_2) \Gamma(0, j_1, t_2, t_2 + j_2) \right| \\
&\leq C \left( \sum_{t_2,t_3,t_4=-\infty}^{\infty} \left| \Gamma(0, t_2, t_3, t_4) \right| \right)^2 \leq C.
\end{aligned}
$$

*Step 5: bound for $K_2$.* Bounding $K_2$ requires additional notation. First set $t_5 = t_1 + j_1$, $t_6 = t_2 + j_1$, $t_7 = t_3 + j_2$ and $t_8 = t_4 + j_2$, and note that $t_5, \dots, t_8$ depend upon $t_1, \dots, t_4$ and $j_1, j_2$ only. For a partition $B = \{B_\ell, \ell = 1, \dots, d_B\}$ of $\{1, \dots, 8\}$, define $d_B = \operatorname{Card} B$, $\Gamma_B(t_1, \dots, t_8) = \prod_{\ell=1}^{d_B} \operatorname{Cum}\left(u_{t_q}, q \in B_\ell\right)$, and recall that $\operatorname{Cum}(u_t) = Eu_t = 0$. Then the largest $d_B$ yielding a non-vanishing $\Gamma_B$ is $d_B = 4$. When $d_B = 4$, $B$ is a pairwise partition of $\{1, \dots, 8\}$ so that $\Gamma_B$ is a product of covariances. Let $B$ be the set of indecomposable partitions of the two-way table

$$
\begin{matrix}
1 & 5 \\
2 & 6 \\
3 & 7 \\
4 & 8
\end{matrix},
$$

see Brillinger (2001, p. 20) for a definition. Then according to Brillinger (2001, Theorem 2.3.2),

$$
\begin{aligned}
&\operatorname{Cum}\left(u_{t_1} u_{t_1+j_1}, u_{t_2} u_{t_2+j_1}, u_{t_3} u_{t_3+j_2}, u_{t_4} u_{t_4+j_2}\right) \\
&= \sum_{B \in \mathcal{B}} \Gamma_B(t_1, \dots, t_8) = \sum_{B \in \mathcal{B}, d_B \leq 3} \Gamma_B(t_1, \dots, t_8) + \sum_{B \in \mathcal{B}, d_B = 4} \Gamma_B(t_1, \dots, t_8).
\end{aligned}
$$

Some properties of partitions in $\mathcal{B}$ are as follows. Call $\{1,5\}$, $\{2,6\}$, $\{3,7\}$ and $\{4,8\}$ fundamental pairs and say that a $B_1$ in a partition $B$ breaks the pair $\{1,5\}$ if $\{1,5\}$ is not a subset of $B_1$. Then partitions $B \in \mathcal{B}$ are such that each $B_\ell \in B$ must break a fundamental pair. Note that fundamental pairs play a symmetric role. Since $t_{q+4} - t_q$ is $j_1$ or $j_2$ with vanishing $k_{j_1}$ or $k_{j_2}$ if $j_1$ or $j_2$ is larger than $p$, the indexes $t_q$ and $t_{q+4}$ of a fundamental pair also play a symmetric role in the computations below. We now discuss the contribution to $K_2$ of partitions of $\{1,\ldots,8\}$ according to the possible values $1,\ldots,4$ of $d_B$. Due to symmetry, we only consider representative partitions for each case.

Under Assumption K and (B.40), the case $d_B = 1$ gives a contribution to $K_2$ bounded by

$$\left| \frac{1}{n^2} \sum_{j_1,j_2=1}^{n-1} k_{j_1} k_{j_2} \sum_{t_1,t_2=1}^{n-j_1} \sum_{t_3,t_4=1}^{n-j_2} \Gamma\left(t_1,\ldots,t_8\right) \right| \leq \frac{C}{n^2} \sum_{t_1,\ldots,t_8=-n}^{n} \left|\Gamma\left(0,t_2-t_1,\ldots,t_8-t_1\right)\right|$$

$$\leq \frac{C}{n} \sum_{t_2',\ldots,t_8'=-\infty}^{\infty} \left|\Gamma\left(0,t_2',\ldots,t_8'\right)\right| \leq \frac{C}{n}.$$

The case $d_B = 2$ corresponds to $\{\operatorname{Card} B_1, \operatorname{Card} B_2\}$ being $\{2,6\}$, $\{3,5\}$ or $\{4,4\}$. These cases are very similar and we limit ourselves to $\{2,6\}$ and $B_1 = \{1,2\}$. The corresponding contribution to $K_2$ is bounded by

$$\left| \frac{1}{n^2} \sum_{j_1,j_2=1}^{n-1} k_{j_1} k_{j_2} \sum_{t_1,t_2=1}^{n-j_1} \sum_{t_3,t_4=1}^{n-j_2} \Gamma_B\left(t_1,\ldots,t_8\right) \right| \leq \frac{C}{n^2} \sum_{t_1,\ldots,t_8=-n}^{n} \left|\Gamma\left(0,t_2-t_1\right)\Gamma\left(t_3-t_1,\ldots,t_8-t_1\right)\right|$$

$$\leq \frac{C}{n} \sum_{t_2',\ldots,t_8'=-n}^{n} \left|\Gamma\left(0,t_2'\right)\Gamma\left(t_3',\ldots,t_8'\right)\right| \leq \frac{C}{n} \sum_{t=-n}^{n} |R_t| \sum_{t_3',\ldots,t_8'=-n}^{n} \left|\Gamma\left(0,t_4'-t_3',\ldots,t_8'-t_3'\right)\right|$$

$$C \sum_{t=-\infty}^{\infty} |R_t| \sum_{t_2,\ldots,t_6=-\infty}^{\infty} \left|\Gamma\left(0,t_2,\ldots,t_6\right)\right| \leq C,$$

by Assumption K and (B.40).

The case $d_B = 3$ corresponds to $\{\operatorname{Card} B_1, \operatorname{Card} B_2, \operatorname{Card} B_3\}$ being $\{2,2,4\}$ or $\{2,3,3\}$. We start with $\operatorname{Card} B_1 = 2$, $\operatorname{Card} B_2 = 2$ and $\operatorname{Card} B_3 = 4$. The discussion concerns the number of fundamental pair broken by $B_3$. Note that the situation where $B_3$ breaks only

3 or 1 fundamental pair is impossible. The case where $B_3$ does not break any fundamental pairs corresponds to partitions that are not indecomposable, so that the only possible cases are those where $B_3$ breaks 4 or 2 fundamental pairs.

- $B_3$ breaks 4 fundamental pairs. Consider $B_3 = \{1,2,3,4\}$, $B_2 = \{5,6\}$ and $B_3 = \{7,8\}$. The corresponding contribution to $K_2$ is bounded by

$$\left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma_B\left(t_1,\dots,t_8\right)\right|$$
$$=\left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma\left(0,t_2-t_1,t_3-t_1,t_4-t_1\right)R_{t_2-t_1}R_{t_4-t_3}\right|$$
$$\leq C\frac{p^2}{n}\sup_j |R_j|^2\sum_{t_2,t_3,t_4=-\infty}^{\infty}\left|\Gamma\left(0,t_2,t_3,t_4\right)\right| \leq C\frac{p^2}{n}$$

  by Assumption K and (B.40).

- $B_3$ breaks 2 fundamental pairs. Take $B_3 = \{1,2,3,5\}$, $B_2 = \{4,6\}$ and $B_1 = \{7,8\}$. The change of variables $t_2 = t_2' + t_1$, $t_3 = t_3' + t_1$ and $t_4 = t_4' + t_3$ shows that contribution to $K_2$ is bounded by

$$\left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma_B\left(t_1,\dots,t_8\right)\right|$$
$$=\left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma\left(0,t_2-t_1,t_3-t_1,j_1\right)R_{t_4-t_2-j_1}R_{t_4-t_3}\right|$$
$$\leq \frac{C}{n}\sum_{j_2=1}^{n-1}K^2(j_2/p)\sum_{t_2',t_3',j_1=-\infty}^{\infty}\left|\Gamma\left(0,t_2',t_3',j_1\right)\right|\sum_{t_4'=-\infty}^{+\infty}\left|R_{t_4'}\right|\times\sup_j |R_j| \leq C\frac{p}{n}.$$

  under Assumption K and (B.40).

We now turn to the case $\mathrm{Card}\, B_3 = \mathrm{Card}\, B_2 = 3$ and $\mathrm{Card}\, B_1 = 2$. Observe that $B_3$ or $B_2$ must break 3 or 1 fundamental pair. The discussion now concerns the fundamental pairs which are simultaneously broken by $B_3$ and $B_2$. Note that $B_3$ and $B_2$ cannot break the same 3 fundamental pairs because if it did, $B_1$ would be given by the remaining fundamental pair in which case $B_1$ cannot communicate with $B_2$ or $B_3$, a fact that would contradict the requirement that the partition $\{B_1, B_2, B_3\}$ is indecomposable.

- $B_3$ and $B_2$ break 3 fundamental pairs, 2 of which are the same. Take $B_3 = \{1, 2, 3\}$, $B_2 = \{4, 5, 6\}$ and $B_1 = \{7, 8\}$. Using change of variables $t_2 = t_1 + t_2'$, $t_3 = t_1 + t_3'$ and $t_4 = t_3 + t_4'$, we can see that under Assumption K and (B.40) the contribution to $K_2$ of this case is bounded by

$$\left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma_B\left(t_1,\ldots,t_8\right)\right|$$
$$= \left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma\left(0,t_2-t_1,t_3-t_1\right)\Gamma\left(0,t_1-t_4+j_1,t_2-t_4+j_1\right)R_{t_4-t_3}\right|$$
$$\leq \frac{C}{n}\sum_{j_1,j_2=1}^{n-1} K^2(j_1/p)K^2(j_2/p)\sup_{t_2,t_3}\left|\Gamma(0,t_2,t_3)\right|\sum_{t_2',t_3'=-\infty}^{\infty}\left|\Gamma\left(0,t_2',t_3'\right)\right|\sum_{t_4'=-\infty}^{+\infty}\left|R_{t_4'}\right| \leq C\frac{p^2}{n}$$

Note that the case where $B_3$ and $B_2$ break 3 fundamental pairs with less than one in common is impossible.

The next case assumes that $B_2$ breaks only 1 fundamental pair, which is also necessarily broken by $B_3$ since $B_2$ must contain the remaining unbroken pair.

- $B_3$ breaks 3 fundamental pairs and $B_2$ breaks only 1 pair. Take $B_3 = \{1, 2, 3\}$, $B_2 = \{4, 5, 8\}$ and $B_3 = \{6, 7\}$ and consider a change of variables $t_2 = t_1 + t_2'$, $t_3 = t_1 + t_3'$ and $t_4 = t_1 + j_1 - t_4'$. Under Assumption K and (B.40), the contribution of this term to $K_2$ is bounded by

$$\left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma_B\left(t_1,\ldots,t_8\right)\right|$$
$$= \left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma\left(0,t_2-t_1,t_3-t_1\right)\Gamma\left(t_1-t_4+j_1,0,j_2\right)R_{t_3-t_2+j_2-j_1}\right|$$
$$\leq \frac{C\sup_j\left|R_j\right|}{n}\sum_{j_1}^{n-1} K^2(j_1/p)\sum_{t_2',t_3'=-\infty}^{\infty}\left|\Gamma(0,t_2',t_3')\right|\sum_{t_4',j_2=-\infty}^{\infty}\left|\Gamma\left(t_4',0,j_2\right)\right| \leq C\frac{p}{n}.$$

- $B_3$ and $B_2$ break only 1 pair. Note that $B_3$ and $B_2$ cannot break the same pair because $B_1$ must be the remaining pair and cannot communicate, so that the partition is not indecomposable. Hence all the partitions in this case are similar to $B_3 = \{1,2,5\}$, $B_2 = \{3,4,8\}$, $B_1 = \{6,7\}$. The change of variable $t_2 = t_1 + t'_2$, $t_3 = -j_2 + t_2 + j_1 + t'_3$ and $t_4 = t_3 - t'_4$ yields a contribution to $K_2$ bounded by

$$\left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma_B\left(t_1,\ldots,t_8\right)\right|$$
$$= \left|\frac{1}{n^2}\sum_{j_1,j_2=1}^{n-1} k_{j_1}k_{j_2}\sum_{t_1,t_2=1}^{n-j_1}\sum_{t_3,t_4=1}^{n-j_2}\Gamma\left(0,t_2-t_1,j_1\right)\Gamma\left(t_3-t_4,0,j_2\right)R_{t_3-t_2+j_2-j_1}\right|$$
$$\leq C\sum_{j_1,t'_2=-\infty}^{\infty}|\Gamma(0,t'_2,j_1)|\sum_{j_2,t'_4=-\infty}^{\infty}|\Gamma(t_4,0,j_2)|\sum_{t'_3=-\infty}^{\infty}\left|R_{t'_3}\right| \leq C.\square$$

## Supplementary material additional references


Brillinger, D.R. (2001). *Time Series Analysis: Data Analysis and Theory* . Holt, Rinehart & Winston, New-York.

Chow, Y.S. and H. Teicher (1988). *Probability Theory. Independence, Interchangeability, Martingales* . Second Edition, Springer.

Li, D. and R.J. Tomkins (1996). Laws of the Iterated Logarithm for Weighted Independent Random Variables. *Statistics and Probability Letters* **27**, 247–254.

Priestley, M.B. (1981). *Spectral Analysis and Time Series.* New York: John Wiley.